EHR Usability Test Report
Safety Enhanced Design 170.315(g)(3)
Computerized Provider Order Entry (CPOE)–Medications(a)1

UCD Process Used: NISTIR 7741:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7741) NIST, Guide to the Processes Approach for Improving the Usability of Electronic Health Records. NIST, 29 Nov. 2010. Web. https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7741.pdf

Report Based on: NISTIR 7742:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7742) NIST, Customized Common Industry Format Template for Electronic Health Record Usability Testing. NIST, 16 Nov. 2010. Web.
https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7742.pdf

Contents

EHR USABILITY TEST REPORT – CPOE: Medications 170.315(a)1

Report based on ISO/IEC 25062:2006 Common Industry Format for Usability Test Reports

Date of Usability Test: December 6, 2018
Date Usability Test was Conducted: December 6, 2018
Date Report was Prepared: April 15, 2019
Report Prepared by: Deborah Ross
Name of System Test Laboratory (STL): Holy Name Medical Center
STL Contact Person, Title and Affiliation: Deborah Ross, Clinical Analyst
STL Phone Number: 201-833-7114
STL Email Address: debross@mail.holyname.org
STL Mailing Address: 718 Teaneck Road, Teaneck, New Jersey 07666

EXECUTIVE SUMMARY

A usability test of CPOE-Medications in WebHIS 2.0 was conducted on December 6, 2018, in Teaneck, New Jersey by Holy Name Medical Center. The purpose of this test was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT).

During the usability test, fifteen (15) healthcare providers matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative tasks.

The study collected performance data on three (3) tasks typically conducted on an EHR:

- **Record Medication Order via CPOE**
- **Change Medication Order via CPOE**
- **Access and record existing medication order**

During the twenty (20) minute summative usability test, each participant was greeted by the administrator and they were instructed that they could withdraw at any time. All participants had varying levels of prior experience with the EHR. The administrator introduced the test, and instructed participants to complete a series of tasks (given one at a time) using the EHRUT. During the testing, the administrator timed the test, and along with the logger(s) recorded user performance data on paper. The administrator did not give the participant assistance in how to complete the task.

The following types of data were collected for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations
- Participant's satisfaction ratings of the system – LIKERT and System Usability Scale (SUS)

All participant data was de-identified – no correspondence could be made from the identity of the participant to the data collected. Various recommended metrics, in accordance with the examples set forth in the NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records were used to evaluate the usability of the EHRUT. The summary data collected for 170.315(a)1 is listed in the **RESULTS** section of this document. The following is the recommended template utilized when calculating the individual results:

Measure 172.315(a)1	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Record medication order	15						
Change medication order	15						
Review medication order	15						

The overall results calculated from the **System Usability Scale (SUS)** provided to each participant scored the subjective satisfaction with the system based on performance with these tasks to be **86 %.** This scale was provided to each participant **only one time at the end of all of the tests**.

In addition to the performance data, qualitative data including both major findings and areas for improvement, based on participant anecdotal feedback will be listed in the RESULTS section as well.

INTRODUCTION

The EHRUT tested for this study was the CPOE-Medications functionality in WebHIS 2.0. Designed to present medical information to healthcare providers in the acute care setting at Holy Name Medical Center, the EHRUT is a clinical information system used by all staff who participate in the care of our patients, on a clinical, financial or record keeping level. The usability testing attempted to represent realistic exercises and conditions.

The purpose of this study was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT). To this end, measures of effectiveness, efficiency, and user satisfaction were captured during the usability testing.

METHOD

PARTICIPANTS

A total of fifteen (15) participants were tested on the EHRUT. Participants in the test were staff nurses, nurses in clinical management, physicians and medical students. Participants work at Holy Name Medical Center in the clinical areas throughout the hospital. Participants did not participate in the requirements gathering, design or development of the EHRUT being tested. Participants were given the

opportunity to have the same orientation and level of training as the actual end users would have received.

Recruited participants had a mix of backgrounds and demographic characteristics conforming to the recruitment screener. The following is a table of participants by characteristics, including demographics, professional experience, computing experience and user needs for assistive technology. Participant names were replaced with Participant IDs so that an individual's data cannot be tied back to individual identities.

	Participant ID	Gender	Age	Education	Occupation/ Role	Professional Experience (mo)	Computer Experience (mo)	Product Experience (mo)	Assistive Technology Needs
1	ID01	Female	30-39	Bachelor's Degree	Staff RN	12	180	12	-
2	ID02	Male	20-29	Bachelor's Degree	Staff RN	60	150	6	-
3	ID03	Female	60-69	Master's Degree	Nursing Director	420	360	120	-
4	ID04	Female	20-29	Bachelor's Degree	Medical Student	0	130	1	-
5	ID05	Female	40-49	Bachelor's Degree	Staff RN	84	240	84	-
6	ID06	Male	30-39	Doctorate Degree	Hospitalist	156	240	36	-
7	ID07	Female	40-49	Associate Degree	Staff RN	240	240	84	-
8	ID08	Male	20-29	Bachelor's Degree	Medical Student	0	180	1	-
9	ID09	Female	40-49	Bachelor's Degree	Staff RN	216	240	72	-
10	ID10	Female	50-59	Master's Degree	Staff RN	420	200	72	-
11	ID11	Male	40-49	Doctorate Degree	Hospitalist	182	192	36	-
12	ID12	Female	60-69	Associate Degree	Operating Room RN	456	120	60	-
13	ID13	Female	30-39	Bachelor's Degree	Staff RN	48	180	24	-
14	ID14	Female	20-29	Bachelor's Degree	Staff RN	36	180	12	-
15	ID15	Female	50-59	Associate Degree	Staff RN	324	240	96	-

Fifteen (15) participants (matching the demographics in the section on Participants) were recruited and fifteen (15) participated in the usability test. Zero (0) participants failed to show for the study.

Participants were scheduled for multiple fifteen (15) and twenty (20) minute sessions with five (5) minutes in between each session for debrief by the administrator(s) and data logger(s), and to reset

systems to proper test conditions where required. A spreadsheet was used to keep track of the participant schedule, and included each participant's demographic characteristics.

STUDY DESIGN

The objective of this test was to uncover areas where the application performed well – that is, effectively, efficiently, and with satisfaction – and areas where the application failed to meet the needs of the participants. The data from this test may serve as a baseline for future tests with an updated version of the same EHR and/or comparison with other EHRs provided the same tasks are used. In short, this testing serves as both a means to record or benchmark current usability, but also to identify areas where improvements must be made.

During the usability test, participants interacted with one EHR. Each participant used the system in the same location, and was provided with the same instructions. The system was evaluated for effectiveness, efficiency and satisfaction as defined by measures collected and analyzed for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

Additional information about the various measures can be found in the section on Usability Metrics.

TASKS

A number of tasks were constructed that would be realistic and representative of the kinds of activities a user might do with this EHR, as follows:

- ***Enter an order for Pantoprazole 40 mg, daily***
- ***Change the above order to Pantoprazole 40 mg, twice a day (BID)***
- ***Review the list of existing medication orders in the CPOE screen***

Tasks were selected based on their frequency of use, criticality of function, and those that may be most troublesome for users. Tasks should always be constructed in light of the study objectives.

PROCEDURES

Upon arrival, participants were greeted by the test administrator. Their identity was verified and matched with their name on the participant schedule. Participants were then assigned an anonymous participant ID.

To ensure that the test ran smoothly, multiple staff members participated in this test – the administrator and several members of the hospital's simulation center, who are experienced test administrators/loggers and frequently participate in these types of exercises in our busy simulation training and testing center.

The administrator moderated the session including administering instructions and tasks. The administrator in conjunction with the data loggers monitored task times, obtained post-task rating data, and took notes on participant comments, in addition to taking notes on task success, path deviations, number and type of errors and comments.

Participants were instructed to perform the tasks (see specific instructions below):

- As quickly as possible making as few errors and deviations as possible.
- Without assistance; administrators were allowed to give immaterial guidance and clarification on tasks, but not instructions on use.
- Without using a think aloud technique.

For each task, the participants were given a written copy of the task. Task timing began once the administrator finished reading the question. The task time was stopped once the participant indicated they had successfully completed the task. Scoring is discussed below.

Participants' demographic information, task success rate, time on task, errors, deviations, verbal responses, and post-test questionnaire were recorded into a spreadsheet.

TEST LOCATION

The test facility included a quiet computer training room with tables and a desktop computer for each participant. To ensure that the environment was comfortable for users, noise levels were kept to a minimum with the ambient temperature within a normal range. All of the safety instruction and evacuation procedures were valid, in place, reviewed with and visible to the participants.

TEST ENVIRONMENT

The EHRUT would typically be used in a hospital (inpatient unit, procedural area, clinic). In this instance, the testing was conducted in the same facility on computers that were set up the same way as the clinical areas of the hospital where the application will be used. The computer being used was an **IBMi Power 7 running I5OS v7.2.** The environment in which the testing was conducted was the WebHIS 2.0 User Acceptance Testing (UAT) environment.

The participants used a mouse and keyboard when interacting with the EHRUT. The following is the configuration of the PCs used both in the test environment as well as throughout the organization in the clinical areas where desktop PCs are used:

Windows 10, dual 2.70GHz IntelCore i5 processor
8GB RAM
1600 x 900 resolution
RGB color format, SDR Color Space
6-bit
60 Hz refresh rate
23 inch monitor

The application was set up by the Holy Name Medical Center Information Systems department '**Desktop Support Division'**, according to the vendor's documentation, describing the system set-up and preparation.

System performance (i.e., response time) on the UAT is typically slower than that of the live environment. This was considered when calculating optimal times into the parameters for testing. Additionally, participants were instructed to not change any of the default system settings (such as control of font size).

TEST FORMS AND TOOLS

During the usability test, various documents and instruments were used, including:

- Participant demographic form
- Description of the test cases
- Objective test data capture form (for each test)
- Likert Scale for each test scenario
- System Usability Scale – one overall for each participant

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 1(a) Enter new medication order in CPOE***
Task	***Enter an order for Pantoprazole 40 mg. daily. Save the order to the order basket but do not place the order.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 1(b) Change medication order in CPOE***
Task	***Change the frequency of Pantoprazole 40 mg. daily to twice daily and then place the order.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 1(c) Review and record medication orders from CPOE screen***
Task	***Review the patient medication list in the CPOE screen and record the medications***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

PARTICIPANT INSTRUCTIONS

The administrator reads the following instructions aloud to the each participant:

> *Thank you for participating in this study. Your input is very important. Our CPOE medication Safety Enhanced Design test session today will last for 20 minutes with a 5 minute break after everyone has completed the exercises. During that time you will use an instance of an electronic health record. I will ask you to complete a few tasks using this system and answer some questions. You should complete the tasks as quickly as possible making as few errors as possible. Please try to complete the tasks on your own following the instructions very closely.*
>
> *Please note that we are not testing you, we are testing the system, therefore if you have difficulty with any aspect of the tasks it likely means that something needs to be improved in the system. Please make note of it if you feel it will help us improve the usability of the system.*
>
> *We will be here in case you need general questions, but we are not able to instruct you or provide help in how to use the application. Overall, we are interested in how easy (or how difficult) this system is to use, what in it would be useful to you, and how we could improve it. Please be honest in your opinions. All of the information that you provide will be kept confidential. Your name will not be associated with any test data or comments at any time. Should you feel it necessary you are able to withdraw at any time during the testing.*

Following the procedural instructions, participants were shown the EHR and as their first task, were given 10 minutes to explore the system and ask general questions. Once this task was complete, the administrator gave the following instructions:

> *For each task, I will read the description to you and say "Begin." At that point, please perform the task and say "Done" once you believe you have successfully completed the task. I would like to request that you not talk aloud or verbalize while you are doing the tasks. I will ask you your impressions about the task once you are done.*

Participants were then given three (3) tasks to complete relating to CPOE – Medications.

USABILITY METRICS

According to the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records*, EHRs should support a process that provides a high level of usability for all users. The goal is for users to interact with the system effectively, efficiently, and with an acceptable level of satisfaction. To this end, metrics for effectiveness, efficiency and user satisfaction were captured during the usability testing. The goals of the test were to assess:

1. Effectiveness of WebHIS 2.0 by measuring participant success rates and errors
2. Efficiency of WebHIS 2.0 by measuring the average task time and path deviations
3. Satisfaction with WebHIS 2.0 by measuring usability, using the System Usability Scale (SUS) overall (at the end of all the testing) and a Likert scale for ease of use for each exercise.

DATA SCORING

The following table details how tasks were scored, errors evaluated, and the time data analyzed.

Measures	Rationale and Scoring
Effectiveness: Task Success	A task was counted as a "Success" if the participant was able to achieve the correct outcome, without assistance, within the time allotted on a per task basis. The total number of successes were calculated for each task and then divided by the total number of times that task was attempted. The results are provided as a percentage. Task times were recorded for successes. Observed task times divided by the optimal time for each task is a measure of optimal efficiency. Optimal task performance time, as benchmarked by expert performance under realistic conditions, is recorded when constructing tasks. Target task times used for task times in the Moderator's Guide must be operationally defined by taking multiple measures of optimal performance and multiplying by some factor [e.g., 1.25] that allows some time buffer because the participants are presumably not trained to expert performance. Thus, if expert, optimal performance on a task was [x] seconds then allotted task time performance was [x * 1.25] seconds. This ratio should be aggregated across tasks and reported with mean and variance scores.
Effectiveness: Task Failures	If the participant abandoned the task, did not reach the correct answer or performed it incorrectly, or reached the end of the allotted time before successful completion, the task was counted as a "Failure." No task times were taken for errors. The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant. On a qualitative level, an enumeration of errors and error types should be collected.
Efficiency: Task Deviations	The participant's path (i.e., steps) through the application was recorded. Deviations occur if the participant, for example, went to a wrong screen, clicked on an incorrect menu item, followed an incorrect link, or interacted incorrectly with an on-screen control. This path was compared to the optimal path. The number of steps in the observed path is divided by the number of optimal steps to provide a ratio of path deviation. It is strongly recommended that task deviations be reported. Optimal paths (i.e., procedural steps) should be recorded when constructing tasks.
Efficiency: Task Time	Each task was timed from when the administrator said "Begin" until the participant said, "Done." If he or she failed to say "Done," the time was stopped when the participant stopped performing the task. Only task times for tasks that were successfully completed were included in the average task time analysis. Average time per task was calculated for each task. Variance measures (standard deviation and standard error) were also calculated.
Satisfaction: Task Rating	Participant's subjective impression of the ease of use of the application was measured by administering both a simple post-task question as well as a post-session questionnaire. After each task, the participant was asked to rate "Overall, this task was:" on a scale of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants. Common convention is that average ratings for systems judged easy to use should be 3.3 or above. To measure participants' confidence in and likeability of the [EHRUT] overall, the testing team administered the System Usability Scale (SUS) post-test questionnaire. Questions included, "I think I would like to use this system frequently," "I thought the system was easy to use," and "I would imagine that most people would learn to use this system very quickly." See full System Usability Score questionnaire in Appendix.

RESULTS

DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above. Participants who failed to follow session and task instructions had their data excluded from the analyses.

The usability testing results for the EHRUT are detailed below. The results should be seen in light of the objectives and goals outlined in the Study Design section. The data should yield actionable results that if corrected, yield material, positive impact on user performance.

Measure 172.315(g)3	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task (a)1	#	Mean (SD)	Deviations (Observed/Optimal) (SD)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Enter med order	15	100%	1.19	80.5 sec	89.5%	1.53	3.79
Change med order	15	100%	1.20	52 sec	86.7%	1.40	3.87
Review med order	15	100%	1.50	34.9 sec	87.3%	1.00	4.00

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 86%. Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80 would be considered above average.

MAJOR FINDINGS

- Major Findings –
  - The CPOE medication ordering screen does not provide the end-user with an easy way to search by route or classification of medication – ie: large volume fluids, injectables, antibiotics, etc.
  - It was not obvious to the untrained user that existing orders could be updated, renewed and modified without placing additional orders (under the 'active medications' list).
  - User comments mentioned the ability to quickly add a STAT medication order.

AREAS FOR IMPROVEMENT

- Areas for Improvement
  - Add a STAT option to the main CPOE screen for rapid ordering of emergency drugs.
  - Differentiate among the various filter buttons. The 'all orders', 'personal favorites', 'order sets', etc. buttons are not obviously differentiated.
  - Highlight the update, renew and modify icons on the CPOE medication list to make them more immediately evident to the user.
  - Provide filters for medication types – ie: large volume fluids, injectables, antibiotics, etc.

EFFECTIVENESS

Overall it was found that the effectiveness of the CPOE – Medication system was very high, based on both participant satisfaction as well as the overall System Usability Scale score. The recommended modifications will make the system more efficient but likely not more effective.

EFFICIENCY

The efficiency of this application based on time –

Enter New Medication order – the optimal time for this task was determined to be 90 seconds. The average participant time as demonstrated during the test was 80.5 seconds.

Change Medication Order – the optimal time for this task was determined to be 60 seconds. The average participant time as demonstrated during the test was 52 seconds.

Review Medication Orders – the optimal time for this task was determined to be 40 seconds. The average participant time as demonstrated during the test was 34.9 seconds.

The efficiency of this application based on path deviations (number of clicks):

Enter New Medication order – the optimal number of clicks for this task was determined to be 8. The average number of clicks demonstrated during the test was 9.53.

Change the Medication Order – the optimal number of clicks for this task was determined to be 7. The average number of clicks demonstrated during this test was 8.40.

Review Medication Order – the optimal number of clicks for this task was determined to be 2. The average number of clicks demonstrated during this test was 3.

SATISFACTION

Utilizing a Likert Scale of 1-5, 5 being 'very easy' to 1 being 'very difficult', the average user rating for each of the Medication CPOE options is as follows:

CPOE – enter medication order	**3.8**
CPOE – change medication order	**3.9**
CPOE - review medication order	**4.0**

Participant comments overall reflected a very high satisfaction rate with the CPOE medication system, as noted in the following quotes collected from the follow up forms:

"Quick and easy to use; it is just like online shopping."
"I had no issues with this test."
"System is simple, easy."

EHR Usability Test Report - 170.315(g)(3)-Safety Enhanced Design Computerized Order Entry-Laboratory(a)2

UCD Process Used: NISTIR 7741:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7741) NIST, Guide to the Processes Approach for Improving the Usability of Electronic Health Records. NIST, 29 Nov. 2010. Web. https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7741.pdf

Report Based on: NISTIR 7742:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7742) NIST, Customized Common Industry Format Template for Electronic Health Record Usability Testing. NIST, 16 Nov. 2010. Web.
https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7742.pdf

Contents

EHR USABILITY TEST REPORT – CPOE: Laboratory Orders 170.315(a)2

Report based on ISO/IEC 25062:2006 Common Industry Format for Usability Test Reports

Date of Usability Test: December 6, 2018
Date Usability Test was Conducted: December 6, 2018
Date Report was Prepared: April 15, 2019
Report Prepared by: Deborah Ross
Name of System Test Laboratory (STL): Holy Name Medical Center
STL Contact Person, Title and Affiliation: Deborah Ross, Clinical Analyst
STL Phone Number: 201-833-7114
STL Email Address: debross@mail.holyname.org
STL Mailing Address: 718 Teaneck Road, Teaneck, New Jersey 07666

EXECUTIVE SUMMARY

A usability test of CPOE-Laboratory Orders in WebHIS 2.0 was conducted on December 6, 2018, in Teaneck, New Jersey by Holy Name Medical Center. The purpose of this test was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT).

During the usability test, fifteen (15) healthcare providers matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative tasks.

The study collected performance data on three (3) tasks typically conducted on an EHR:

- **Enter Laboratory Test Order via CPOE**
- **Change Laboratory Test Order via CPOE**
- **Access and record existing laboratory result**

During the twenty (20) minute summative usability test, each participant was greeted by the administrator and they were instructed that they could withdraw at any time. All participants had varying levels of prior experience with the EHR. The administrator introduced the test, and instructed participants to complete a series of tasks (given one at a time) using the EHRUT. During the testing, the administrator timed the test, and along with the logger(s) recorded user performance data on paper. The administrator did not give the participant assistance in how to complete the task.

The following types of data were collected for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations
- Participant's satisfaction ratings of the system – LIKERT and System Usability Scale (SUS)

All participant data was de-identified – no correspondence could be made from the identity of the participant to the data collected. Various recommended metrics, in accordance with the examples set forth in the NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records were used to evaluate the usability of the EHRUT. The summary data collected for **170.315(a)2** is listed in the **RESULTS** section of this document. The following is the recommended template utilized when calculating the individual results:

Measure 172.315(a)2	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Enter lab order	15						
Change lab order	15						
Review lab results	15						

The overall results calculated from the **System Usability Scale (SUS)** provided to each participant scored the subjective satisfaction with the system based on performance with these tasks to be **86 %.** This scale was provided to each participant **only one time at the end of all of the tests**.

In addition to the performance data, qualitative data including both major findings and areas for improvement, based on participant anecdotal feedback will be listed in the RESULTS section as well.

INTRODUCTION

The EHRUT tested for this study were tasks associated with CPOE-Laboratory Order functionality in WebHIS 2.0. Designed to present medical information to healthcare providers in the acute care setting at Holy Name Medical Center, the EHRUT is a clinical information system used by all staff who participate in the care of our patients, on a clinical, financial or record keeping level. The usability testing attempted to represent realistic exercises and conditions.

The purpose of this study was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT). To this end, measures of effectiveness, efficiency, and user satisfaction were captured during the usability testing.

METHOD

PARTICIPANTS

A total of fifteen (15) participants were tested on the EHRUT. Participants in the test were staff nurses, nurses in clinical management, physicians and medical students. Participants work at Holy Name Medical Center in the clinical areas throughout the hospital. Participants did not participate in the requirements gathering, design or development of the EHRUT being tested. Participants were given the opportunity to have the same orientation and level of training as the actual end users would have received.

Recruited participants had a mix of backgrounds and demographic characteristics conforming to the recruitment screener. The following is a table of participants by characteristics, including demographics, professional experience, computing experience and user needs for assistive technology. Participant

names were replaced with Participant IDs so that an individual's data cannot be tied back to individual identities.

	Participant ID	Gender	Age	Education	Occupation/ Role	Professional Experience (mo)	Computer Experience (mo)	Product Experience (mo)	Assistive Technology Needs
1	ID01	Female	30-39	Bachelor's Degree	Staff RN	12	180	12	-
2	ID02	Male	20-29	Bachelor's Degree	Staff RN	60	150	6	-
3	ID03	Female	60-69	Master's Degree	Nursing Director	420	360	120	-
4	ID04	Female	20-29	Bachelor's Degree	Medical Student	0	130	1	-
5	ID05	Female	40-49	Bachelor's Degree	Staff RN	84	240	84	-
6	ID06	Male	30-39	Doctorate Degree	Hospitalist	156	240	36	-
7	ID07	Female	40-49	Associate Degree	Staff RN	240	240	84	-
8	ID08	Male	20-29	Bachelor's Degree	Medical Student	0	180	1	-
9	ID09	Female	40-49	Bachelor's Degree	Staff RN	216	240	72	-
10	ID10	Female	50-59	Master's Degree	Staff RN	420	200	72	-
11	ID11	Male	40-49	Doctorate Degree	Hospitalist	182	192	36	-
12	ID12	Female	60-69	Associate Degree	Operating Room RN	456	120	60	-
13	ID13	Female	30-39	Bachelor's Degree	Staff RN	48	180	24	-
14	ID14	Female	20-29	Bachelor's Degree	Staff RN	36	180	12	-
15	ID15	Female	50-59	Associate Degree	Staff RN	324	240	96	-

Fifteen (15) participants (matching the demographics in the section on Participants) were recruited and fifteen (15) participated in the usability test. Zero (0) participants failed to show for the study.

Participants were scheduled for multiple fifteen (15) and twenty (20) minute sessions in the course of all the testing, with five (5) minutes in between each session for debrief by the administrator(s) and data logger(s), and to reset systems to proper test conditions where required. A spreadsheet was used to keep track of the participant schedule and included each participant's demographic characteristics.

STUDY DESIGN

The objective of this test was to uncover areas where the application performed well – that is, effectively, efficiently, and with satisfaction – and areas where the application failed to meet the needs of the participants. The data from this test may serve as a baseline for future tests with an updated

version of the same EHR and/or comparison with other EHRs provided the same tasks are used. In short, this testing serves as both a means to record or benchmark current usability, and a means to identify areas where improvements should be made.

During the usability test, participants interacted with one EHR. Each participant used the system in the same location, and was provided with the same instructions. The system was evaluated for effectiveness, efficiency and satisfaction as defined by measures collected and analyzed for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

Additional information about the various measures can be found in the section on Usability Metrics.

TASKS

A number of tasks were constructed that would be realistic and representative of the kinds of activities a user might do with this EHR, as follows:

- ***Enter an order for a routine morning CBC***
- ***Change the above order to a STAT order***
- ***Identify and record the patient's most recent LDL results***

Tasks were selected based on their frequency of use and criticality of function. Tasks should always be constructed in light of the study objectives.

PROCEDURES

Upon arrival, participants were greeted by the test administrator. Their identity was verified and matched with their name on the participant schedule. Participants were then assigned an anonymous participant ID.

To ensure that the test ran smoothly, multiple staff members participated in this test – the administrator and several members of the hospital's simulation center, who are experienced test administrators/loggers and frequently participate in these types of exercises in our busy simulation training and testing center.

The administrator moderated the session including administering instructions and tasks. The administrator in conjunction with the data loggers monitored task times, obtained post-task rating data, and took notes on participant comments, in addition to taking notes on task success, path deviations, number and type of errors and comments.

Participants were instructed to perform the tasks (see specific instructions below):

- As quickly as possible making as few errors and deviations as possible.
- Without assistance; administrators were allowed to give immaterial guidance and clarification on tasks, but not instructions on use.
- Without using a think aloud technique.

For each task, the participants were given a written copy of the task. Task timing began once the administrator finished reading the question. The task time was stopped once the participant indicated they had successfully completed the task. Scoring is discussed below.

Participants' demographic information, task success rate, time on task, errors, deviations, verbal responses, and post-test questionnaire were recorded into a spreadsheet.

TEST LOCATION

The test facility included a quiet computer training room with tables and a desktop computer for each participant. To ensure that the environment was comfortable for users, noise levels were kept to a minimum with the ambient temperature within a normal range. All of the safety instruction and evacuation procedures were valid, in place, reviewed with and visible to the participants.

TEST ENVIRONMENT

The EHRUT would typically be used in a hospital (inpatient unit, procedural area, clinic). In this instance, the testing was conducted in the same facility on computers that were set up the same way as the clinical areas of the hospital where the application will be used. The computer being used was an **IBMi Power 7 running I5OS v7.2.** The environment in which the testing was conducted was the WebHIS 2.0 User Acceptance Testing (UAT) environment.

The participants used a mouse and keyboard when interacting with the EHRUT. The following is the configuration of the PCs used both in the test environment as well as throughout the organization in the clinical areas where desktop PCs are used:

Windows 10, dual 2.70GHz IntelCore i5 processor
8GB RAM
1600 x 900 resolution
RGB color format, SDR Color Space
6-bit
60 Hz refresh rate
23 inch monitor

The application was set up by the Holy Name Medical Center Information Systems department **Desktop Support Division** according to the vendor's documentation, describing the system set-up and preparation.

System performance (i.e., response time) on the UAT is typically slower than that of the live environment. This was considered when calculating optimal times into the parameters for testing. Additionally, participants were instructed to not change any of the default system settings (such as control of font size).

TEST FORMS AND TOOLS

During the usability test, various documents and instruments were used, including:

- Participant demographic form
- Description of the test cases
- Objective test data capture form (for each test)
- Likert Scale for each test scenario
- System Usability Scale – one overall for each participant

Objective Test Data Forms for 170.315(a)2

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 2(a) Enter new Laboratory order in CPOE***
Task	***Enter an order for Complete Blood Count (CBC) for early AM. Save the order to the order basket but do not place the order.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 2(b) Change the Laboratory order in CPOE***
Task	***Change the Complete Blood Count (CBC) order from early AM to STAT.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 2(c) Review patient's laboratory orders.***
Task	***In patient's laboratory orders, identify most recent Lipid Profile and record the LDL.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

PARTICIPANT INSTRUCTIONS

The administrator reads the following instructions aloud to the each participant:

> *Thank you for participating in this study. Your input is very important. Our CPOE Laboratory Orders Enhanced Design test session today will last for 20 minutes with a 5 minute break after everyone has completed the exercises. During that time you will use an instance of an electronic health record. I will ask you to complete a few tasks using this system and answer some questions. You should complete the tasks as quickly as possible making as few errors as possible. Please try to complete the tasks on your own following the instructions very closely.*
>
> *Please note that we are not testing you, we are testing the system, therefore if you have difficulty with any aspect of the tasks it likely means that something needs to be improved in the system. Please make note of it if you feel it will help us improve the usability of the system.*
>
> *We will be here in case you have general questions, but we are not able to instruct you or provide help in how to use the application. Overall, we are interested in how easy (or how difficult) this system is to use, what in it would be useful to you, and how we could improve it. Please be honest in your opinions. All of the information that you provide will be kept confidential. Your name will not be associated with any test data or comments at any time. Should you feel it necessary you are able to withdraw at any time during the testing.*

Following the procedural instructions, participants were shown the EHR and as their first task, were given 10 minutes to explore the system and ask general questions. Once this task was complete, the administrator gave the following instructions:

> *For each task, I will read the description to you and say "Begin." At that point, please perform the task and say "Done" once you believe you have successfully completed the task. I would like to request that you not talk aloud or verbalize while you are doing the tasks. I will ask you your impressions about the task once you are done.*

Participants were then given three (3) tasks to complete relating to ***CPOE – Laboratory Orders.***

USABILITY METRICS

According to the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records*, EHRs should support a process that provides a high level of usability for all users. The goal is for users to interact with the system effectively, efficiently, and with an acceptable level of satisfaction. To this end, metrics for effectiveness, efficiency and user satisfaction were captured during the usability testing. The goals of the test were to assess:

1. Effectiveness of WebHIS 2.0 by measuring participant success rates and errors
2. Efficiency of WebHIS 2.0 by measuring the average task time and path deviations
3. Satisfaction with WebHIS 2.0 by measuring usability, using the System Usability Scale (SUS) overall (at the end of all the testing) and a Likert scale for ease of use for each exercise.

DATA SCORING

The following table details how tasks were scored, errors evaluated, and the time data analyzed.

Measures	Rationale and Scoring
Effectiveness: Task Success	A task was counted as a "Success" if the participant was able to achieve the correct outcome, without assistance, within the time allotted on a per task basis. The total number of successes were calculated for each task and then divided by the total number of times that task was attempted. The results are provided as a percentage. Task times were recorded for successes. Observed task times divided by the optimal time for each task is a measure of optimal efficiency. Optimal task performance time, as benchmarked by expert performance under realistic conditions, is recorded when constructing tasks. Target task times used for task times in the Moderator's Guide must be operationally defined by taking multiple measures of optimal performance and multiplying by some factor [e.g., 1.25] that allows some time buffer because the participants are presumably not trained to expert performance. Thus, if expert, optimal performance on a task was [x] seconds then allotted task time performance was [x * 1.25] seconds. This ratio should be aggregated across tasks and reported with mean and variance scores.
Effectiveness: Task Failures	If the participant abandoned the task, did not reach the correct answer or performed it incorrectly, or reached the end of the allotted time before successful completion, the task was counted as a "Failure." No task times were taken for errors. The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant. On a qualitative level, an enumeration of errors and error types should be collected.
Efficiency: Task Deviations	The participant's path (i.e., steps) through the application was recorded. Deviations occur if the participant, for example, went to a wrong screen, clicked on an incorrect menu item, followed an incorrect link, or interacted incorrectly with an on-screen control. This path was compared to the optimal path. The number of steps in the observed path is divided by the number of optimal steps to provide a ratio of path deviation. It is strongly recommended that task deviations be reported. Optimal paths (i.e., procedural steps) should be recorded when constructing tasks.
Efficiency: Task Time	Each task was timed from when the administrator said "Begin" until the participant said, "Done." If he or she failed to say "Done," the time was stopped when the participant stopped performing the task. Only task times for tasks that were successfully completed were included in the average task time analysis. Average time per task was calculated for each task. Variance measures (standard deviation and standard error) were also calculated.
Satisfaction: Task Rating	Participant's subjective impression of the ease of use of the application was measured by administering both a simple post-task question as well as a post-session questionnaire. After each task, the participant was asked to rate "Overall, this task was:" on a scale of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants. Common convention is that average ratings for systems judged easy to use should be 3.3 or above. To measure participants' confidence in and likeability of the [EHRUT] overall, the testing team administered the System Usability Scale (SUS) post-test questionnaire. Questions included, "I think I would like to use this system frequently," "I thought the system was easy to use," and "I would imagine that most people would learn to use this system very quickly." See full System Usability Score questionnaire in Appendix.

RESULTS

DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above. Participants who failed to follow session and task instructions had their data excluded from the analyses.

The usability testing results for the EHRUT are detailed below. The results should be seen in light of the objectives and goals outlined in the Study Design section. The data should yield actionable results that if corrected, yield material, positive impact on user performance.

Measure 172.315(g)3	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task (a)2	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean
Enter lab order	15	100%	1.29	50.7 sec	84.4%	1.73	4.00
Change order	15	100%	1.20	44.5 sec	89.0%	0.6	4.13
Review results	15	100%	1.10	36.5 sec	81.0%	0.40	4.07

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 86%. Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80 would be considered above average.

MAJOR FINDINGS

- **Major Findings –**
  - **The CPOE Laboratory order screen does not provide the end-user with an easy way to search by laboratory department – blood bank, microbiology, etc.**
  - **It was not obvious to the untrained user that an order already in the order basket could be changed as opposed to deleting the order and entering a new one with the newest criteria.**
  - **There is no STAT lab order entry or 'quick list' to select from.**

AREAS FOR IMPROVEMENT

- **Areas for Improvement**
  - **Add a STAT option to the main CPOE screen for rapid ordering of emergency drugs.**
  - **Differentiate among the various filter buttons. The 'all orders', 'personal favorites', 'order sets', etc. buttons are not obviously differentiated.**
  - **Include the ability to filter laboratory orders by individual department for easier searching.**

EFFECTIVENESS

Overall it was found that the effectiveness of the **CPOE – Laboratory Order** system was very high, based on both participant satisfaction as well as the overall System Usability Scale score. The recommended modifications will make the system more efficient but likely not more effective.

EFFICIENCY

The efficiency of this application based on time

Enter new laboratory order – the optimal time for this task was determined to be 60 seconds. The average time as demonstrated during the test was 50.7 seconds.

Modify the laboratory order – the optimal time for this task was determined to be 50 seconds. The average time as demonstrated during the test was 44.5 seconds.

Review results of specified test – the optimal time for this task was determined to be 45 seconds. The average time as demonstrated during the test was 36.5 (36.47) seconds.

Efficiency of this application based on path deviations (number of clicks):

Enter new laboratory order – the optimal number of clicks for this task was determined to be 6. The average number of clicks demonstrated during the test was 7.7.

Modify the laboratory order – the optimal number of clicks for this task was determined to be 3. The average number of clicks demonstrated during this test was 3.6

Review results of specified test – the optimal number of clicks for this task was determined to be 4. The average number of clicks demonstrated during this test was 4.4.

SATISFACTION

Utilizing a Likert Scale of 1-5, 5 being 'very easy' to 1 being 'very difficult', the average user rating for each of the Laboratory test cases is as follows:

CPOE – enter laboratory order	4.0
CPOE – change laboratory order	4.36
CPOE – retrieve laboratory results	4.13

Participant comments overall reflected a very high satisfaction rate with the CPOE laboratory system, as noted in the following quotes collected from the follow up forms:

"Quick and easy to use; it is just like online shopping."

"Finding results was very easy. I even checked out the trending."

"Overall, the system is easy, but it would be faster if we could sort out the lab departments."

"Please add a 'STAT' quick list for faster entry of stat orders."

EHR Usability Test Report –
170.315(g)(3)– Safety Enhanced Design
Computerized Provider Order Entry – Radiology Orders (a)3

UCD Process Used: NISTIR 7741:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7741) NIST, Guide to the Processes Approach for Improving the Usability of Electronic Health Records. NIST, 29 Nov. 2010. Web. https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7741.pdf

Report Based on: NISTIR 7742:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7742) NIST, Customized Common Industry Format Template for Electronic Health Record Usability Testing. NIST, 16 Nov. 2010. Web.
https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7742.pdf

Contents

EHR USABILITY TEST REPORT – CPOE: Radiology Orders 170.315(a)3

Report based on ISO/IEC 25062:2006 Common Industry Format for Usability Test Reports

Date of Usability Test: December 6, 2018
Date Usability Test was Conducted: December 6, 2018
Date Report was Prepared: April 15, 2019
Report Prepared by: Deborah Ross
Name of System Test Laboratory (STL): Holy Name Medical Center
STL Contact Person, Title and Affiliation: Deborah Ross, Clinical Analyst
STL Phone Number: 201-833-7114
STL Email Address: debross@mail.holyname.org
STL Mailing Address: 718 Teaneck Road, Teaneck, New Jersey 07666

EXECUTIVE SUMMARY

A usability test of CPOE-Radiology Orders in WebHIS 2.0 was conducted on December 6, 2018, in Teaneck, New Jersey by Holy Name Medical Center. The purpose of this test was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT).

During the usability test, fifteen (15) healthcare providers matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative tasks.

The study collected performance data on three (3) tasks typically conducted on an EHR:

- Enter a Radiology Order via CPOE
- Change the Radiology Order via CPOE
- Access and review an existing Radiology result

During the twenty (20) minute summative usability test, each participant was greeted by the administrator, and they were instructed that they could withdraw at any time. All participants had varying levels of prior experience with the EHR. The administrator introduced the test, and instructed participants to complete a series of tasks (given one at a time) using the EHRUT. During the testing, the administrator timed the test, and along with the logger(s) recorded user performance data on paper. The administrator did not give the participant assistance in how to complete the task.

The following types of data were collected for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations
- Participant's satisfaction ratings of the system – LIKERT and System Usability Scale (SUS)

All participant data was de-identified – no correspondence could be made from the identity of the participant to the data collected. Various recommended metrics, in accordance with the examples set forth in the NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records were used to evaluate the usability of the EHRUT. The summary data collected for ***170.315(a)3*** is listed in the **RESULTS** section of this document. The following is the recommended template utilized when calculating the individual results:

Measure 172.315(a)3	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task (a)3	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Enter Radiology order	15						
Change Radiology order	15						
Review Radiology Report	15						

The overall results calculated from the **System Usability Scale (SUS)** provided to each participant scored the subjective satisfaction with the system based on performance with these tasks to be **86 %.** This scale was provided to each participant **only one time at the end of all of the tests**.

In addition to the performance data, qualitative data including both major findings and areas for improvement, based on participant anecdotal feedback will be listed in the RESULTS section as well.

INTRODUCTION

The EHRUT tested for this study was the CPOE-Radiology Order functionality in WebHIS 2.0. Designed to present medical information to healthcare providers in the acute care setting at Holy Name Medical Center, the EHRUT is a clinical information system used by all staff who participate in the care of our patients, on a clinical, financial or record keeping level. The usability testing attempted to represent realistic exercises and conditions.

The purpose of this study was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT). To this end, measures of effectiveness, efficiency, and user satisfaction were captured during the usability testing.

METHOD

PARTICIPANTS

A total of fifteen (15) participants were tested on the EHRUT. Participants in the test were staff nurses, nurses in clinical management, physicians and medical students. Participants work at Holy Name Medical Center in the clinical areas throughout the hospital. Participants did not participate in the requirements gathering, design or development of the EHRUT being tested. Participants were given the

opportunity to have the same orientation and level of training as the actual end users would have received.

Recruited participants had a mix of backgrounds and demographic characteristics conforming to the recruitment screener. The following is a table of participants by characteristics, including demographics, professional experience, computing experience and user needs for assistive technology. Participant names were replaced with Participant IDs so that an individual's data cannot be tied back to individual identities.

	Participant ID	Gender	Age	Education	Occupation/ Role	Professional Experience (mo)	Computer Experience (mo)	Product Experience (mo)	Assistive Technology Needs
1	ID01	Female	30-39	Bachelor's Degree	Staff RN	12	180	12	-
2	ID02	Male	20-29	Bachelor's Degree	Staff RN	60	150	6	-
3	ID03	Female	60-69	Master's Degree	Nursing Director	420	360	120	-
4	ID04	Female	20-29	Bachelor's Degree	Medical Student	0	130	1	-
5	ID05	Female	40-49	Bachelor's Degree	Staff RN	84	240	84	-
6	ID06	Male	30-39	Doctorate Degree	Hospitalist	156	240	36	-
7	ID07	Female	40-49	Associate Degree	Staff RN	240	240	84	-
8	ID08	Male	20-29	Bachelor's Degree	Medical Student	0	180	1	-
9	ID09	Female	40-49	Bachelor's Degree	Staff RN	216	240	72	-
10	ID10	Female	50-59	Master's Degree	Staff RN	420	200	72	-
11	ID11	Male	40-49	Doctorate Degree	Hospitalist	182	192	36	-
12	ID12	Female	60-69	Associate Degree	Operating Room RN	456	120	60	-
13	ID13	Female	30-39	Bachelor's Degree	Staff RN	48	180	24	-
14	ID14	Female	20-29	Bachelor's Degree	Staff RN	36	180	12	-
15	ID15	Female	50-59	Associate Degree	Staff RN	324	240	96	-

Fifteen (15) participants (matching the demographics in the section on Participants) were recruited and fifteen (15) participated in the usability test. Zero (0) participants failed to show for the study.

Participants were scheduled for multiple fifteen (15) and twenty (20) minute sessions with five (5) minutes in between each session for debrief by the administrator(s) and data logger(s), and to reset

systems to proper test conditions where required A spreadsheet was used to keep track of the participant schedule, and included each participant's demographic characteristics.

STUDY DESIGN

Overall the objective of this test was to uncover areas where the application performed well – that is, effectively, efficiently, and with satisfaction – and areas where the application failed to meet the needs of the participants. The data from this test may serve as a baseline for future tests with an updated version of the same EHR and/or comparison with other EHRs provided the same tasks are used. In short, this testing serves as both a means to record or benchmark current usability, but also to identify areas where improvements must be made.

During the usability test, participants interacted with one EHR. Each participant used the system in the same location, and was provided with the same instructions. The system was evaluated for effectiveness, efficiency and satisfaction as defined by measures collected and analyzed for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

Additional information about the various measures can be found in the section on Usability Metrics.

TASKS

A number of tasks were constructed that would be realistic and representative of the kinds of activities a user might do with this EHR, as follows:

- ***Enter an order for a Left Ankle x-ray due to fracture and don't place order***
- ***Change the above order to a Right Ankle and add the reason 'dislocation'***
- ***Identify and review the patient's most recent portable chest x-ray result***

Tasks were selected based on their frequency of use and criticality of function. Tasks should always be constructed in light of the study objectives.

PROCEDURES

Upon arrival, participants were greeted by the test administrator. Their identity was verified and matched with their name on the participant schedule. Participants were then assigned an anonymous participant ID.

To ensure that the test ran smoothly, multiple staff members participated in this test – the administrator and several members of the hospital's simulation center, who are experienced test administrators/loggers and frequently participate in these types of exercises in our busy simulation training and testing center.

The administrator moderated the session including administering instructions and tasks. The administrator in conjunction with the data loggers monitored task times, obtained post-task rating data, and took notes on participant comments, in addition to taking notes on task success, path deviations, number and type of errors and comments.

Participants were instructed to perform the tasks (see specific instructions below):

- As quickly as possible making as few errors and deviations as possible.
- Without assistance; administrators were allowed to give immaterial guidance and clarification on tasks, but not instructions on use.
- Without using a think aloud technique.

For each task, the participants were given a written copy of the task. Task timing began once the administrator finished reading the question. The task time was stopped once the participant indicated they had successfully completed the task. Scoring is discussed below.

Participants' demographic information, task success rate, time on task, errors, deviations, verbal responses, and post-test questionnaire were recorded into a spreadsheet.

TEST LOCATION

The test facility included a quiet computer training room with tables and a desktop computer for each participant. To ensure that the environment was comfortable for users, noise levels were kept to a minimum with the ambient temperature within a normal range. All of the safety instruction and evacuation procedures were valid, in place, reviewed with and visible to the participants.

TEST ENVIRONMENT

The EHRUT would typically be used in a hospital (inpatient unit, procedural area, clinic). In this instance, the testing was conducted in the same facility on computers that were set up the same way as the clinical areas of the hospital where the application will be used. The computer being used was an **IBMi Power 7 running I5OS v7.2.** The environment in which the testing was conducted was the WebHIS 2.0 User Acceptance Testing (UAT) environment.

The participants used a mouse and keyboard when interacting with the EHRUT. The following is the configuration of the PCs used both in the test environment as well as throughout the organization in the clinical areas where desktop PCs are used:

Windows 10, dual 2.70GHz IntelCore i5 processor
8GB RAM
1600 x 900 resolution
RGB color format, SDR Color Space
6-bit
60 Hz refresh rate
23 inch monitor

The application was set up by the Holy Name Medical Center Information Systems department '**Desktop Support Division'**, according to the vendor's documentation, describing the system set-up and preparation.

System performance (i.e., response time) on the UAT is typically slower than that of the live environment. This was considered when calculating optimal times into the parameters for testing. Additionally, participants were instructed to not change any of the default system settings (such as control of font size).

TEST FORMS AND TOOLS

During the usability test, various documents and instruments were used, including:

- Participant demographic form
- Description of the test cases
- Objective test data capture form (for each test)
- Likert Scale for each test scenario
- System Usability Scale – one overall for each participant

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 3(a) Enter a Radiology order in CPOE***
Task	***Enter an order for a Left Ankle x-ray due to fracture. Do not place the order.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 3(b) Modify a Radiology order in CPOE***
Task	***Change the above order to Right ankle and add the reason 'dislocation'. Place the order.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 3(c) Review and record x-ray results***
Task	***Identify your patient's most recent portable chest x-ray order and review the impression.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

PARTICIPANT INSTRUCTIONS

The administrator reads the following instructions aloud to the each participant:

> *Thank you for participating in this study. Your input is very important. Our* CPOE Radiology Orders Enhanced Design test session today will last for 20 minutes with a 5 minute break *after everyone has completed the exercises. During that time you will use an instance of an electronic health record. I will ask you to complete a few tasks using this system and answer some questions. You should complete the tasks as quickly as possible making as few errors as possible. Please try to complete the tasks on your own following the instructions very closely.*
>
> *Please note that we are not testing you, we are testing the system, therefore if you have difficulty with any aspect of the tasks it likely means that something needs to be improved in the system. Please make note of it if you feel it will help us improve the usability of the system.*
>
> *We will be here in case you need general questions, but we are not able to instruct you or provide help in how to use the application. Overall, we are interested in how easy (or how difficult) this system is to use, what in it would be useful to you, and how we could improve it. Please be honest in your opinions. All of the information that you provide will be kept confidential. Your name will not be associated with any test data or comments at any time. Should you feel it necessary you are able to withdraw at any time during the testing.*

Following the procedural instructions, participants were shown the EHR and as their first task, were given 10 minutes to explore the system and ask general questions. Once this task was complete, the administrator gave the following instructions:

> *For each task, I will read the description to you and say "Begin." At that point, please perform the task and say "Done" once you believe you have successfully completed the task. I would like to request that you not talk aloud or verbalize while you are doing the tasks. I will ask you your impressions about the task once you are done.*

Participants were then given three (3) tasks to complete relating to CPOE – Radiology Orders.

USABILITY METRICS

According to the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records*, EHRs should support a process that provides a high level of usability for all users. The goal is for users to interact with the system effectively, efficiently, and with an acceptable level of satisfaction. To this end, metrics for effectiveness, efficiency and user satisfaction were captured during the usability testing. The goals of the test were to assess:

1. Effectiveness of WebHIS 2.0 by measuring participant success rates and errors
2. Efficiency of WebHIS 2.0 by measuring the average task time and path deviations
3. Satisfaction with WebHIS 2.0 by measuring usability, using the System Usability Scale (SUS) overall (at the end of all the testing) and a Likert scale for ease of use for each exercise.

DATA SCORING

The following table details how tasks were scored, errors evaluated, and the time data analyzed.

Measures	Rationale and Scoring
Effectiveness: Task Success	A task was counted as a "Success" if the participant was able to achieve the correct outcome, without assistance, within the time allotted on a per task basis. The total number of successes were calculated for each task and then divided by the total number of times that task was attempted. The results are provided as a percentage. Task times were recorded for successes. Observed task times divided by the optimal time for each task is a measure of optimal efficiency. Optimal task performance time, as benchmarked by expert performance under realistic conditions, is recorded when constructing tasks. Target task times used for task times in the Moderator's Guide must be operationally defined by taking multiple measures of optimal performance and multiplying by some factor [e.g., 1.25] that allows some time buffer because the participants are presumably not trained to expert performance. Thus, if expert, optimal performance on a task was [x] seconds then allotted task time performance was [x * 1.25] seconds. This ratio should be aggregated across tasks and reported with mean and variance scores.
Effectiveness: Task Failures	If the participant abandoned the task, did not reach the correct answer or performed it incorrectly, or reached the end of the allotted time before successful completion, the task was counted as a "Failure." No task times were taken for errors. The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant. On a qualitative level, an enumeration of errors and error types should be collected.
Efficiency: Task Deviations	The participant's path (i.e., steps) through the application was recorded. Deviations occur if the participant, for example, went to a wrong screen, clicked on an incorrect menu item, followed an incorrect link, or interacted incorrectly with an on-screen control. This path was compared to the optimal path. The number of steps in the observed path is divided by the number of optimal steps to provide a ratio of path deviation. It is strongly recommended that task deviations be reported. Optimal paths (i.e., procedural steps) should be recorded when constructing tasks.
Efficiency: Task Time	Each task was timed from when the administrator said "Begin" until the participant said, "Done." If he or she failed to say "Done," the time was stopped when the participant stopped performing the task. Only task times for tasks that were successfully completed were included in the average task time analysis. Average time per task was calculated for each task. Variance measures (standard deviation and standard error) were also calculated.
Satisfaction: Task Rating	Participant's subjective impression of the ease of use of the application was measured by administering both a simple post-task question as well as a post-session questionnaire. After each task, the participant was asked to rate "Overall, this task was:" on a scale of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants. Common convention is that average ratings for systems judged easy to use should be 3.3 or above. To measure participants' confidence in and likeability of the [EHRUT] overall, the testing team administered the System Usability Scale (SUS) post-test questionnaire. Questions included, "I think I would like to use this system frequently," "I thought the system was easy to use," and "I would imagine that most people would learn to use this system very quickly." See full System Usability Score questionnaire in Appendix.

RESULTS

DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above. Participants who failed to follow session and task instructions had their data excluded from the analyses.

The usability testing results for the EHRUT are detailed below. The results should be seen in light of the objectives and goals outlined in the Study Design section. The data should yield actionable results that if corrected, yield material, positive impact on user performance.

Measure 172.315(a)3	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task	#	Mean (SD)	Deviations Observed/Optimal	Mean (SD)	Deviations Observed/Optimal	Mean (SD)	Mean
Enter Radiology order	15	100%	1.16	53.13 sec	75.9%	1.4	4.00
Change Radiology order	15	100%	1.14	37.7 sec	83.7%	0.87	4.07
Review Radiology Report	15	100%	1.17	36.2 sec	80.4%	0.33	4.13

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 86%. Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80 would be considered above average.

MAJOR FINDINGS

- Major Findings –
  - The CPOE Radiology ordering screen does not provide the end-user with an easy way to search by diagnostic department – CT Scans, MRI, General Diagnostics, etc.
  - It was not obvious to the untrained user that an order sitting in the order basket could be changed as opposed to deleting the order and entering a new one with the newest criteria.
  - There is no STAT radiology order entry or 'quick list' to select from.
  - It was not obvious at the outset that there was only a single order for each body part that involves 'laterality'. Some users stated searching using 'left' in the search field rather than ankle, with no results.

AREAS FOR IMPROVEMENT

- Areas for Improvement
  - Add a STAT option to the main CPOE screen for rapid ordering of emergency tests.

- Differentiate among the various filter buttons - 'all orders', 'personal favorites', 'order sets', etc. buttons are not obviously differentiated. Will consider using color to differentiate.
- Include the ability to filter Radiology orders by individual department for easier searching – ie: CT Scan, MRI, etc.

EFFECTIVENESS

Overall it was found that the effectiveness of the CPOE – Radiology Order system was very high, based on both participant satisfaction as well as the overall System Usability Scale score. The recommended modifications will make the system more efficient but likely not more effective.

EFFICIENCY

The efficiency of this application based on time -

Enter new radiology order – the optimal time for this task was determined to be 70 seconds. The average time as demonstrated during the test was 53 seconds. The efficiency rate for time was calculated to be 88.3%.
Modify the radiology order – the optimal time for this task was determined to be 45 seconds. The average time as demonstrated during the test was 38 seconds (37.6). The efficiency rate calculated for time was 83.7%.
Review radiology report – the optimal time for this task was determined to be 45 seconds. The average time as demonstrated during the test was 36 seconds. The efficiency rate calculated for time was 80.4%.

Efficiency of this application based on path deviations (number of clicks) -

Enter new radiology order – the optimal number of clicks for this task was determined to be 9. The average number of clicks demonstrated during the test was 10.4. The efficiency demonstrated by the number of clicks for this task was calculated to be 88.3%.
Change the radiology order – the optimal number of clicks for this task was determined to be 6. The average number of clicks demonstrated during this test was 6.9. The efficiency demonstrated by the number of clicks for this task was calculated to be 89.9%
Review radiology report – The optimal number of clicks for this task was determined to be 2. The average number of clicks demonstrated during this 2.3. The efficiency demonstrated by the number of clicks for this task was calculated to be 90%.

SATISFACTION

Utilizing a Likert Scale of 1-5, 5 being 'very easy' to 1 being 'very difficult', the average user rating for each of the Radiology tasks is as follows:

CPOE – Enter Radiology order	4.00
CPOE – Change Radiology order	4.07
CPOE – Review Radiology report	4.13

Participant comments overall reflected a very high satisfaction rate with the CPOE medication system, as noted in the following quotes collected from the follow up forms:

"Quick and easy to use; it is just like online shopping."

"Finding correct reports was relatively easy. Is there a way to search by term? That would be easier." (note – there is a way to do this)

"Overall, the system is easy, but it would be faster if we could sort out the Radiology departments when placing orders."

EHR Usability Test Report
170.315(g)(3) Safety Enhanced Design
CPOE Drug/Drug & Drug/Allergy Interaction Checking (a)4

UCD Process Used: NISTIR 7741:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7741) NIST, Guide to the Processes Approach for Improving the Usability of Electronic Health Records. NIST, 29 Nov. 2010. Web. https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7741.pdf

Report Based on: NISTIR 7742:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7742) NIST, Customized Common Industry Format Template for Electronic Health Record Usability Testing. NIST, 16 Nov. 2010. Web.
https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7742.pdf

Contents

EHR USABILITY TEST REPORT – Drug/Drug & Drug/Allergy Interaction Checking

Report based on ISO/IEC 25062:2006 Common Industry Format for Usability Test Reports

Date of Usability Test: December 6, 2018
Date Usability Test was Conducted: December 6, 2018
Date Report was Prepared: April 15, 2019
Report Prepared by: Deborah Ross
Name of System Test Laboratory (STL): Holy Name Medical Center
STL Contact Person, Title and Affiliation: Deborah Ross, Clinical Analyst
STL Phone Number: 201-833-7114
STL Email Address: debross@mail.holyname.org
STL Mailing Address: 718 Teaneck Road, Teaneck, New Jersey 07666

EXECUTIVE SUMMARY

A usability test of CPOE-Drug/Drug and Drug/Allergy Interaction Checking in WebHIS 2.0 was conducted on December 6, 2018, in Teaneck, New Jersey by Holy Name Medical Center. The purpose of this test was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT).

During the usability test, fifteen (15) healthcare providers matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative tasks.

The study collected performance data on two (2) tasks typically conducted on an EHR:

- Enter an order for Penicillin 500 mg p.o. and verify that an allergy alert launches
- Enter an order for Ondansetron 2 mg. IV and for Amiodarone 200 mg p.o. and note the interaction alert

During the twenty (20) minute summative usability test, each participant was greeted by the administrator and they were instructed that they could withdraw at any time. All participants had varying levels of prior experience with the EHR. The administrator introduced the test, and instructed participants to complete a series of tasks (given one at a time) using the EHRUT. During the testing, the administrator timed the test, and along with the logger(s) recorded user performance data on paper. The administrator did not give the participant assistance in how to complete the task.

The following types of data were collected for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations
- Participant's satisfaction ratings of the system – LIKERT and System Usability Scale (SUS)

All participant data was de-identified – no correspondence could be made from the identity of the participant to the data collected. Various recommended metrics, in accordance with the examples set forth in the NIST Guide to the Processes Approach for Improving the Usability of Electronic Health

Records were used to evaluate the usability of the EHRUT. The summary data collected for 170.315(a)4 is listed in the **RESULTS** section of this document. The following is the recommended template utilized when calculating the individual results:

Measure 172.315g(3)	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task (a)4	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Drug/Drug Interaction	15						
Drug/Allergy Alert	15						

The overall results calculated from the **System Usability Scale (SUS)** provided to each participant scored the subjective satisfaction with the system based on performance with these tasks to be **86 %.** This scale was provided to each participant **only one time at the end of all of the tests**.

In addition to the performance data, qualitative data including both major findings and areas for improvement, based on participant anecdotal feedback will be listed in the RESULTS section as well.

INTRODUCTION

The EHRUT tested for this study utilized the CPOE-Medication ordering functionality in WebHIS 2.0, but included the added features of **Medication Allergy Alert,** based on the patient's documented allergies, as well as **Drug/Drug interactions** based on evidence-based databases utilized by the EHRUT. Designed to present medical information to healthcare providers in the acute care setting at Holy Name Medical Center, the EHRUT is a clinical information system used by all staff who participate in the care of our patients, on a clinical, financial or record keeping level. The usability testing attempted to represent realistic exercises and conditions.

The purpose of this study was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT). To this end, measures of effectiveness, efficiency, and user satisfaction were captured during the usability testing.

METHOD

PARTICIPANTS

A total of fifteen (15) participants were tested on the EHRUT. Participants in the test were staff nurses, nurses in clinical management, physicians and medical students. Participants work at Holy Name Medical Center in the clinical areas throughout the hospital. Participants did not participate in the requirements gathering, design or development of the EHRUT being tested. Participants were given the opportunity to have the same orientation and level of training as the actual end users would have received.

Recruited participants had a mix of backgrounds and demographic characteristics conforming to the recruitment screener. The following is a table of participants by characteristics, including demographics, professional experience, computing experience and user needs for assistive technology. Participant names were replaced with Participant IDs so that an individual's data cannot be tied back to individual identities.

	Participant ID	Gender	Age	Education	Occupation/ Role	Professional Experience (mo)	Computer Experience (mo)	Product Experience (mo)	Assistive Technology Needs
1	ID01	Female	30-39	Bachelor's Degree	Staff RN	12	180	12	-
2	ID02	Male	20-29	Bachelor's Degree	Staff RN	60	150	6	-
3	ID03	Female	60-69	Master's Degree	Nursing Director	420	360	120	-
4	ID04	Female	20-29	Bachelor's Degree	Medical Student	0	130	1	-
5	ID05	Female	40-49	Bachelor's Degree	Staff RN	84	240	84	-
6	ID06	Male	30-39	Doctorate Degree	Hospitalist	156	240	36	-
7	ID07	Female	40-49	Associate Degree	Staff RN	240	240	84	-
8	ID08	Male	20-29	Bachelor's Degree	Medical Student	0	180	1	-
9	ID09	Female	40-49	Bachelor's Degree	Staff RN	216	240	72	-
10	ID10	Female	50-59	Master's Degree	Staff RN	420	200	72	-
11	ID11	Male	40-49	Doctorate Degree	Hospitalist	182	192	36	-
12	ID12	Female	60-69	Associate Degree	Operating Room RN	456	120	60	-
13	ID13	Female	30-39	Bachelor's Degree	Staff RN	48	180	24	-
14	ID14	Female	20-29	Bachelor's Degree	Staff RN	36	180	12	-
15	ID15	Female	50-59	Associate Degree	Staff RN	324	240	96	-

Fifteen (15) participants (matching the demographics in the section on Participants) were recruited and fifteen (15) participated in the usability test. Zero (0) participants failed to show for the study.

Participants were scheduled for multiple fifteen (15) and twenty (20) minute sessions with five (5) minutes in between each session for debrief by the administrator(s) and data logger(s), and to reset systems to proper test conditions where required A spreadsheet was used to keep track of the participant schedule, and included each participant's demographic characteristics.

STUDY DESIGN

The objective of this test was to uncover areas where the application performed well – that is, effectively, efficiently, and with satisfaction – and areas where the application failed to meet the needs of the participants. The data from this test may serve as a baseline for future tests with an updated version of the same EHR and/or comparison with other EHRs provided the same tasks are used. In short, this testing serves as both a means to record or benchmark current usability, but also to identify areas where improvements must be made.

During the usability test, participants interacted with one EHR. Each participant used the system in the same location, and was provided with the same instructions. The system was evaluated for effectiveness, efficiency and satisfaction as defined by measures collected and analyzed for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

Additional information about the various measures can be found in the section on Usability Metrics.

TASKS

A number of tasks were constructed that would be realistic and representative of the kinds of activities a user might do with this EHR, as follows:

- ***For your patient with a documented penicillin allergy, enter a CPOE order for Penicillin 500 mg p.o. four (4) times a day. Note whether or not an allergy alert appears at the time of ordering.***
- ***Enter a CPOE order for Amiodarone 200 mg. p.o. daily into the basket. Then enter an order for Ondansetron 2mg every six (6) hours. Note whether or not an interaction notification appears and note whether or not you are able to modify this alert.***

Tasks were selected based on their frequency of use, criticality of function, and those that may be most troublesome for users. Tasks should always be constructed in light of the study objectives.

PROCEDURES

Upon arrival, participants were greeted by the test administrator. Their identity was verified and matched with their name on the participant schedule. Participants were then assigned an anonymous participant ID.

To ensure that the test ran smoothly, multiple staff members participated in this test – the administrator and several members of the hospital's simulation center, who are experienced test administrators/loggers and frequently participate in these types of exercises in our busy simulation training and testing center.

The administrator moderated the session including administering instructions and tasks. The administrator in conjunction with the data loggers monitored task times, obtained post-task rating data,

and took notes on participant comments, in addition to taking notes on task success, path deviations, number and type of errors and comments.

Participants were instructed to perform the tasks (see specific instructions below):

- As quickly as possible making as few errors and deviations as possible.
- Without assistance; administrators were allowed to give immaterial guidance and clarification on tasks, but not instructions on use.
- Without using a think aloud technique.

For each task, the participants were given a written copy of the task. Task timing began once the administrator finished reading the question. The task time was stopped once the participant indicated they had successfully completed the task. Scoring is discussed below.

Participants' demographic information, task success rate, time on task, errors, deviations, verbal responses, and post-test questionnaire were recorded into a spreadsheet.

TEST LOCATION

The test facility included a quiet computer training room with tables and a desktop computer for each participant. To ensure that the environment was comfortable for users, noise levels were kept to a minimum with the ambient temperature within a normal range. All of the safety instruction and evacuation procedures were valid, in place, reviewed with and visible to the participants.

TEST ENVIRONMENT

The EHRUT would typically be used in a hospital (inpatient unit, procedural area, clinic). In this instance, the testing was conducted in the same facility on computers that were set up the same way as the clinical areas of the hospital where the application will be used. The computer being used was an **IBMi Power 7 running I5OS v7.2.** The environment in which the testing was conducted was the WebHIS 2.0 User Acceptance Testing (UAT) environment.

The participants used a mouse and keyboard when interacting with the EHRUT. The following is the configuration of the PCs used both in the test environment as well as throughout the organization in the clinical areas where desktop PCs are used:

Windows 10, dual 2.70GHz IntelCore i5 processor
8GB RAM
1600 x 900 resolution
RGB color format, SDR Color Space
6-bit
60 Hz refresh rate
23 inch monitor

The application was set up by the Holy Name Medical Center Information Systems department **Desktop Support Division** according to the vendor's documentation, describing the system set-up and preparation.

System performance (i.e., response time) on the UAT is typically slower than that of the live environment. This was considered when calculating optimal times into the parameters for testing. Additionally, participants were instructed to not change any of the default system settings (such as control of font size).

TEST FORMS AND TOOLS

During the usability test, various documents and instruments were used, including:

- Participant demographic form
- Description of the test cases
- Objective test data capture form (for each test)
- Likert Scale for each test scenario
- System Usability Scale – one overall for each participant

Sample Objective Test Data Forms:

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 4(a) Drug/Allergy Interactions***
Task	***For your patient with a documented Penicillin allergy, enter an order for Penicillin 500 mg. p.o. four times a day. Note the allergy alert that pops up.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 4(a) Drug/drug interaction***
Task	***Enter the following medication orders for your patient: Amiodarone 200 mg p.o. B.I.D. and Ondansetron 2 mg IV every 12 hours. Note the severe interaction warning. Note that the warning cannot be altered by the end user.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

PARTICIPANT INSTRUCTIONS

The administrator reads the following instructions aloud to the each participant:

> *Thank you for participating in this study. Your input is very important. Our Drug/Drug and Drug/Allergy interaction checking test session today will last for 20 minutes with a 5 minute break after everyone has completed the exercises. During that time you will use an instance of an electronic health record. I will ask you to complete a few tasks using this system and answer some questions. You should complete the tasks as quickly as possible making as few errors as possible. Please try to complete the tasks on your own following the instructions very closely.*
>
> *Please note that we are not testing you, we are testing the system, therefore if you have difficulty with any aspect of the tasks it likely means that something needs to be improved in the system. Please make note of it if you feel it will help us improve the usability of the system.*
>
> *We will be here in case you need general questions, but we are not able to instruct you or provide help in how to use the application. Overall, we are interested in how easy (or how difficult) this system is to use, what in it would be useful to you, and how we could improve it. Please be honest in your opinions. All of the information that you provide will be kept confidential. Your name will not be associated with any test data or comments at any time. Should you feel it necessary you are able to withdraw at any time during the testing.*

Following the procedural instructions, participants were shown the EHR and as their first task, were given 10 minutes to explore the system and ask general questions. Once this task was complete, the administrator gave the following instructions:

> *For each task, I will read the description to you and say "Begin." At that point, please perform the task and say "Done" once you believe you have successfully completed the task. I would like to request that you not talk aloud or verbalize while you are doing the tasks. I will ask you your impressions about the task once you are done.*

Participants were then given two (2) tasks to complete relating to CPOE Medication Allergy and Drug/Drug Interaction Alerts.

USABILITY METRICS

According to the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records*, EHRs should support a process that provides a high level of usability for all users. The goal is for users to interact with the system effectively, efficiently, and with an acceptable level of satisfaction. To this end, metrics for effectiveness, efficiency and user satisfaction were captured during the usability testing. The goals of the test were to assess:

1. Effectiveness of WebHIS 2.0 by measuring participant success rates and errors
2. Efficiency of WebHIS 2.0 by measuring the average task time and path deviations
3. Satisfaction with WebHIS 2.0 by measuring usability, using the System Usability Scale (SUS) overall (at the end of all the testing) and a Likert scale for ease of use for each exercise.

DATA SCORING

The following table details how tasks were scored, errors evaluated, and the time data analyzed.

Measures	Rationale and Scoring
Effectiveness: Task Success	A task was counted as a "Success" if the participant was able to achieve the correct outcome, without assistance, within the time allotted on a per task basis. The total number of successes were calculated for each task and then divided by the total number of times that task was attempted. The results are provided as a percentage. Task times were recorded for successes. Observed task times divided by the optimal time for each task is a measure of optimal efficiency. Optimal task performance time, as benchmarked by expert performance under realistic conditions, is recorded when constructing tasks. Target task times used for task times in the Moderator's Guide must be operationally defined by taking multiple measures of optimal performance and multiplying by some factor [e.g., 1.25] that allows some time buffer because the participants are presumably not trained to expert performance. Thus, if expert, optimal performance on a task was [x] seconds then allotted task time performance was [x * 1.25] seconds. This ratio should be aggregated across tasks and reported with mean and variance scores.
Effectiveness: Task Failures	If the participant abandoned the task, did not reach the correct answer or performed it incorrectly, or reached the end of the allotted time before successful completion, the task was counted as a "Failure." No task times were taken for errors. The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant. On a qualitative level, an enumeration of errors and error types should be collected.
Efficiency: Task Deviations	The participant's path (i.e., steps) through the application was recorded. Deviations occur if the participant, for example, went to a wrong screen, clicked on an incorrect menu item, followed an incorrect link, or interacted incorrectly with an on-screen control. This path was compared to the optimal path. The number of steps in the observed path is divided by the number of optimal steps to provide a ratio of path deviation. It is strongly recommended that task deviations be reported. Optimal paths (i.e., procedural steps) should be recorded when constructing tasks.
Efficiency: Task Time	Each task was timed from when the administrator said "Begin" until the participant said, "Done." If he or she failed to say "Done," the time was stopped when the participant stopped performing the task. Only task times for tasks that were successfully completed were included in the average task time analysis. Average time per task was calculated for each task. Variance measures (standard deviation and standard error) were also calculated.
Satisfaction: Task Rating	Participant's subjective impression of the ease of use of the application was measured by administering both a simple post-task question as well as a post-session questionnaire. After each task, the participant was asked to rate "Overall, this task was:" on a scale of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants. Common convention is that average ratings for systems judged easy to use should be 3.3 or above. To measure participants' confidence in and likeability of the [EHRUT] overall, the testing team administered the System Usability Scale (SUS) post-test questionnaire. Questions included, "I think I would like to use this system frequently," "I thought the system was easy to use," and "I would imagine that most people would learn to use this system very quickly." See full System Usability Score questionnaire in Appendix.

RESULTS

DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above. Participants who failed to follow session and task instructions had their data excluded from the analyses.

The usability testing results for the EHRUT are detailed below. The results should be seen in light of the objectives and goals outlined in the Study Design section. The data should yield actionable results that if corrected, yield material, positive impact on user performance.

Measure 172.315(g)3	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task (a)4	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Note drug/ allergy alert		100%	1.27	50.73 sec	84.6%	1.60	4.00
Note drug/drug interaction alert		100%	1.07	62.00 sec	82.7%	0.60	4.40

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 86%. Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80 would be considered above average.

MAJOR FINDINGS

- Major Findings –
  - The drug/allergy alert is reliable and provides an added safety measure to the system, eliminating reliance on manually checking patient allergies.
  - The drug/drug interaction program and subsequent end user alert provides evidence-based clinical decision support for the prescriber or nurse transcribing the telephone order. This is a very positive system attribute.
  - Displaying the 'conflict' alert (for both of these tests) within the order links the user response to the alerts and maintains them at the order level. This is a positive feature.
  - Displaying the level of severity, onset and link to the evidence for drug/drug interactions was an added bonus beyond simply getting notification of the interaction.
  - Safety of this application is further enhanced by the fact that the alerts can only be added/modified by system administrators. End users conducting this test were required to respond to the alerts, and could not modify/delete them in any way.

AREAS FOR IMPROVEMENT

- There were no recommendations from the participants regarding improvement for either of these programs. The notifications, both for patient allergies and for drug/drug interactions were clearly presented and had good usability.

EFFECTIVENESS

Overall it was found that the effectiveness of these alerts satisfied their intended purpose. The risk of ordering a medication to which the patient has reported an allergy is reduced significantly to almost 0. Users do not have to manually check an allergy list and are reassured that the checking will occur 'behind the scenes' providing an alert when required.

Whether a provider is entering an order or an order is being transcribed as a telephone order by a nurse, the risk of placing an order for medications which have documented interactions is very high. This knowledge is best maintained in a knowledge-base that works seamlessly with the CPOE system. In addition to checking and reporting the drug/drug interactions, the severity level, onset of symptoms and links to the evidence are embedded in the alert as well. This is a very valuable and effective Clinical Decision Support tool.

EFFICIENCY

The efficiency of this application based on time –

Drug/Allergy Alert – the optimal time in which to complete this task was determined to be 60 seconds. The average time as demonstrated during the test was 50.7 seconds.

Drug/Drug Interaction Alert – the optimal time for this task was determined to be 75 seconds. The average time as demonstrated during the test was 62 seconds.

Efficiency of this application based on path deviations (number of clicks) -

Drug/Allergy Alert – the optimal number of clicks for this task was determined to be 6. The average number of clicks demonstrated during the test was 7.6.

Drug/Drug Interaction Alert – the optimal number of clicks for this task was determined to be 9. The average number of clicks demonstrated during this test was 9.6

SATISFACTION

Utilizing a Likert Scale of 1-5, 5 being 'very easy' to 1 being 'very difficult', the average user rating for each of the Drug Alert options is as follows:

Drug/Allergy Alert	4.0
Drug/Drug Interaction Alert	4.4

Participant comments overall reflected a high satisfaction rate with the Allergy and Interaction alerts, as noted in the following quotes collected from the follow up forms:

> "It is very reassuring how the CPOE program alerts us to allergies."
> "It is a very good program, and informative."
> "Easy to understand – helpful pop-ups."
> "Straight forward – no difficulties."
> "Knowing the severity of the reactions was great extra information."
> "Very helpful to show drug reaction alerts."

EHR Usability Test Report -170.315(g)(3)

170.315(a)5 – Demographics

UCD Process Used: NISTIR 7741:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7741) NIST, Guide to the Processes Approach for Improving the Usability of Electronic Health Records. NIST, 29 Nov. 2010. Web. https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7741.pdf

Report Based on: NISTIR 7742:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7742) NIST, Customized Common Industry Format Template for Electronic Health Record Usability Testing. NIST, 16 Nov. 2010. Web.
https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7742.pdf

Contents

EHR USABILITY TEST REPORT – Demographics 170.315(a)5

Report based on ISO/IEC 25062:2006 Common Industry Format for Usability Test Reports

Date of Usability Test: December 6, 2018
Date Usability Test was Conducted: December 6, 2018
Date Report was Prepared: April 15, 2019
Report Prepared by: Deborah Ross
Name of System Test Laboratory (STL): Holy Name Medical Center
STL Contact Person, Title and Affiliation: Deborah Ross, Clinical Analyst
STL Phone Number: 201-833-7114
STL Email Address: debross@mail.holyname.org
STL Mailing Address: 718 Teaneck Road, Teaneck, New Jersey 07666

EXECUTIVE SUMMARY

A usability test of the Demographic information entry and editing application in WebHIS 2.0 was conducted on December 6, 2018, in Teaneck, New Jersey by Holy Name Medical Center. The purpose of this test was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT).

During the usability test, fifteen (15) healthcare providers matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative tasks.

The study collected performance data and user reactions to multiple Demographic editing functions built into the EHRUT.

- Enter and change patient's ethnicity and sexual preference
- Enter and change patient's smoking status
- Record patient date and cause of death

During the twenty (20) minute summative usability test, each participant was greeted by the administrator and they were instructed that they could withdraw at any time. All participants had varying levels of prior experience with the EHR. The administrator introduced the test, and instructed participants to complete a series of tasks (given one at a time) using the EHRUT. During the testing, the administrator timed the test, and along with the logger(s) recorded user performance data on paper. The administrator did not give the participant assistance in how to complete the task.

The following types of data were collected for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations
- Participant's satisfaction ratings of the system – LIKERT and System Usability Scale (SUS)

All participant data was de-identified – no correspondence could be made from the identity of the participant to the data collected. Various recommended metrics, in accordance with the examples set forth in the NIST Guide to the Processes Approach for Improving the Usability of Electronic Health

Records were used to evaluate the usability of the EHRUT. The summary data collected for *170.315(a)5* is listed in the RESULTS section of this document. The following is the recommended template utilized when calculating the individual results:

Measure 172.315(a)5	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Enter & Change Smoking Status	15						
Enter and change patient Ethnicity and Sexual Preference	15						
Enter Cause and Date of Death	15						

The overall results calculated from the **System Usability Scale (SUS)** provided to each participant scored the subjective satisfaction with the system based on performance with these tasks to be **86 %.** This scale was provided to each participant **only one time at the end of all of the tests**.

In addition to the performance data, qualitative data including both major findings and areas for improvement, based on participant anecdotal feedback will be listed in the RESULTS section as well.

INTRODUCTION

The EHRUT tested for this study was the Demographic entry and modification screen in WebHIS 2.0. Designed to present medical information to healthcare providers in the acute care setting at Holy Name Medical Center, the EHRUT is a clinical information system used by all staff who participate in the care of our patients, on a clinical, financial or record keeping level. The usability testing attempted to represent realistic exercises and conditions.

The purpose of this study was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT). To this end, measures of effectiveness, efficiency, and user satisfaction were captured during the usability testing.

METHOD

PARTICIPANTS

A total of fifteen (15) participants were tested on the EHRUT. Participants in the test were staff nurses, nurses in clinical management, physicians and medical students. Participants work at Holy Name Medical Center in the clinical areas throughout the hospital. Participants did not participate in the requirements gathering, design or development of the EHRUT being tested. Participants were given the opportunity to have the same orientation and level of training as the actual end users would have received.

Recruited participants had a mix of backgrounds and demographic characteristics conforming to the recruitment screener. The following is a table of participants by characteristics, including demographics, professional experience, computing experience and user needs for assistive technology. Participant names were replaced with Participant IDs so that an individual's data cannot be tied back to individual identities.

	Participant ID	Gender	Age	Education	Occupation/ Role	Professional Experience (mo)	Computer Experience (mo)	Product Experience (mo)	Assistive Technology Needs
1	ID01	Female	30-39	Bachelor's Degree	Staff RN	12	180	12	-
2	ID02	Male	20-29	Bachelor's Degree	Staff RN	60	150	6	-
3	ID03	Female	60-69	Master's Degree	Nursing Director	420	360	120	-
4	ID04	Female	20-29	Bachelor's Degree	Medical Student	0	130	1	-
5	ID05	Female	40-49	Bachelor's Degree	Staff RN	84	240	84	-
6	ID06	Male	30-39	Doctorate Degree	Hospitalist	156	240	36	-
7	ID07	Female	40-49	Associate Degree	Staff RN	240	240	84	-
8	ID08	Male	20-29	Bachelor's Degree	Medical Student	0	180	1	-
9	ID09	Female	40-49	Bachelor's Degree	Staff RN	216	240	72	-
10	ID10	Female	50-59	Master's Degree	Staff RN	420	200	72	-
11	ID11	Male	40-49	Doctorate Degree	Hospitalist	182	192	36	-
12	ID12	Female	60-69	Associate Degree	Operating Room RN	456	120	60	-
13	ID13	Female	30-39	Bachelor's Degree	Staff RN	48	180	24	-
14	ID14	Female	20-29	Bachelor's Degree	Staff RN	36	180	12	-
15	ID15	Female	50-59	Associate Degree	Staff RN	324	240	96	-

Fifteen (15) participants (matching the demographics in the section on Participants) were recruited and fifteen (15) participated in the usability test. Zero (0) participants failed to show for the study.

Participants were scheduled for multiple fifteen (15) and twenty (20) minute sessions with five (5) minutes in between each session for debrief by the administrator(s) and data logger(s), and to reset systems to proper test conditions where required A spreadsheet was used to keep track of the participant schedule, and included each participant's demographic characteristics.

STUDY DESIGN

The objective of this test was to uncover areas where the application performed well – that is, effectively, efficiently, and with satisfaction – and areas where the application failed to meet the needs of the participants. The data from this test may serve as a baseline for future tests with an updated version of the same EHR and/or comparison with other EHRs provided the same tasks are used. In short, this testing serves as both a means to record or benchmark current usability, but also to identify areas where improvements must be made.

During the usability test, participants interacted with one EHR. Each participant used the system in the same location, and was provided with the same instructions. The system was evaluated for effectiveness, efficiency and satisfaction as defined by measures collected and analyzed for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

Additional information about the various measures can be found in the section on Usability Metrics.

TASKS

A number of tasks were constructed that would be realistic and representative of the kinds of activities a user might do with this EHR, as follows:

- ***Enter patient's ethnicity as Cuban and change to Puerto Rican; enter patient's sexual preference as 'straight' or 'heterosexual' and change to 'choose not to disclose".***
- ***Enter patient's smoking status as 'current everyday smoker' then change to 'former smoker'***
- ***Enter patient's cause of death as 'cardio-pulmonary' arrest and enter current date as date of death.***

Tasks were selected based on their frequency of use, criticality of function, and those that may be most troublesome for users. Tasks should always be constructed in light of the study objectives.

PROCEDURES

Upon arrival, participants were greeted by the test administrator. Their identity was verified and matched with their name on the participant schedule. Participants were then assigned an anonymous participant ID.

To ensure that the test ran smoothly, multiple staff members participated in this test – the administrator and several members of the hospital's simulation center, who are experienced test administrators/loggers and frequently participate in these types of exercises in our busy simulation training and testing center.

The administrator moderated the session including administering instructions and tasks. The administrator in conjunction with the data loggers monitored task times, obtained post-task rating data, and took notes on participant comments, in addition to taking notes on task success, path deviations, number and type of errors and comments.

Participants were instructed to perform the tasks (see specific instructions below):

- As quickly as possible making as few errors and deviations as possible.
- Without assistance; administrators were allowed to give immaterial guidance and clarification on tasks, but not instructions on use.
- Without using a think aloud technique.

For each task, the participants were given a written copy of the task. Task timing began once the administrator finished reading the question. The task time was stopped once the participant indicated they had successfully completed the task. Scoring is discussed below.

Participants' demographic information, task success rate, time on task, errors, deviations, verbal responses, and post-test questionnaire were recorded into a spreadsheet.

TEST LOCATION

The test facility included a quiet computer training room with tables and a desktop computer for each participant. To ensure that the environment was comfortable for users, noise levels were kept to a minimum with the ambient temperature within a normal range. All of the safety instruction and evacuation procedures were valid, in place, reviewed with and visible to the participants.

TEST ENVIRONMENT

The EHRUT would typically be used in a hospital (inpatient unit, procedural area, clinic). In this instance, the testing was conducted in the same facility on computers that were set up the same way as the clinical areas of the hospital where the application will be used. The computer being used was an **IBMi Power 7 running I5OS v7.2.** The environment in which the testing was conducted was the WebHIS 2.0 User Acceptance Testing (UAT) environment.

The participants used a mouse and keyboard when interacting with the EHRUT. The following is the configuration of the PCs used both in the test environment as well as throughout the organization in the clinical areas where desktop PCs are used:

Windows 10, dual 2.70GHz IntelCore i5 processor
8GB RAM
1600 x 900 resolution
RGB color format, SDR Color Space
6-bit
60 Hz refresh rate

23 inch monitor

The application was set up by the Holy Name Medical Center Information Systems department **Desktop Support Division**, according to the vendor's documentation, describing the system set-up and preparation.

System performance (i.e., response time) on the UAT is typically slower than that of the live environment. This was considered when calculating optimal times into the parameters for testing. Additionally, participants were instructed to not change any of the default system settings (such as control of font size).

TEST FORMS AND TOOLS

During the usability test, various documents and instruments were used, including:

- Participant demographic form
- Description of the test cases
- Objective test data capture form (for each test)
- Likert Scale for each test scenario
- System Usability Scale – one overall for each participant

Sample Objective Test Data Forms:

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task (a)5 Demographic Information – Adding and modifying patient Ethnicity and Sexual Preference***
Task	***In patient demographics, enter your patient's ethnicity as Cuban and save. Edit ethnicity to Puerto Rican. Enter patient's sexual preference as 'straight or heterosexual' and save, then change to 'choose not to disclose'.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task (a)5 Demographic Information – Smoking Status***
Task	***Enter the patient's smoking status as 'current everyday smoker' and save it. Enter the demographic tab and modify the smoking status to 'former smoker' and save. Review the current smoking status.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task (a)5 Demographic Information – Cause and date of death***
Task	***In patient demographics, enter your patient's cause of death to 'cardio-pulmonary arrest' with the date of today.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

PARTICIPANT INSTRUCTIONS

The administrator reads the following instructions aloud to the each participant:

> *Thank you for participating in this study. Your input is very important. The Demographics entry and editing test session today will last for 20 minutes with a 5 minute break after everyone has completed the exercises. During that time you will use an instance of an electronic health record. I will ask you to complete a few tasks using this system and answer some questions. You should complete the tasks as quickly as possible making as few errors as possible. Please try to complete the tasks on your own following the instructions very closely.*
>
> *Please note that we are not testing you, we are testing the system, therefore if you have difficulty with any aspect of the tasks it likely means that something needs to be improved in the system. Please make note of it if you feel it will help us improve the usability of the system.*
>
> *We will be here in case you need general questions, but we are not able to instruct you or provide help in how to use the application. Overall, we are interested in how easy (or how difficult) this system is to use, what in it would be useful to you, and how we could improve it. Please be honest in your opinions. All of the information that you provide will be kept confidential. Your name will not be associated with any test data or comments at any time. Should you feel it necessary you are able to withdraw at any time during the testing.*

Following the procedural instructions, participants were shown the EHR and as their first task, were given 10 minutes to explore the system and ask general questions. Once this task was complete, the administrator gave the following instructions:

> *For each task, I will read the description to you and say "Begin." At that point, please perform the task and say "Done" once you believe you have successfully completed the task. I would like to request that you not talk aloud or verbalize while you are doing the tasks. I will ask you your impressions about the task once you are done.*

Participants were then given three (3) tasks to complete relating to Demographics.

USABILITY METRICS

According to the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records*, EHRs should support a process that provides a high level of usability for all users. The goal is for users to interact with the system effectively, efficiently, and with an acceptable level of satisfaction. To this end, metrics for effectiveness, efficiency and user satisfaction were captured during the usability testing. The goals of the test were to assess:

1. Effectiveness of WebHIS 2.0 by measuring participant success rates and errors
2. Efficiency of WebHIS 2.0 by measuring the average task time and path deviations
3. Satisfaction with WebHIS 2.0 by measuring usability, using the System Usability Scale (SUS) overall (at the end of all the testing) and a Likert scale for ease of use for each exercise.

DATA SCORING

The following table details how tasks were scored, errors evaluated, and the time data analyzed.

Measures	Rationale and Scoring
Effectiveness: Task Success	A task was counted as a "Success" if the participant was able to achieve the correct outcome, without assistance, within the time allotted on a per task basis. The total number of successes were calculated for each task and then divided by the total number of times that task was attempted. The results are provided as a percentage. Task times were recorded for successes. Observed task times divided by the optimal time for each task is a measure of optimal efficiency. Optimal task performance time, as benchmarked by expert performance under realistic conditions, is recorded when constructing tasks. Target task times used for task times in the Moderator's Guide must be operationally defined by taking multiple measures of optimal performance and multiplying by some factor [e.g., 1.25] that allows some time buffer because the participants are presumably not trained to expert performance. Thus, if expert, optimal performance on a task was [x] seconds then allotted task time performance was [x * 1.25] seconds. This ratio should be aggregated across tasks and reported with mean and variance scores.
Effectiveness: Task Failures	If the participant abandoned the task, did not reach the correct answer or performed it incorrectly, or reached the end of the allotted time before successful completion, the task was counted as a "Failure." No task times were taken for errors. The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant. On a qualitative level, an enumeration of errors and error types should be collected.
Efficiency: Task Deviations	The participant's path (i.e., steps) through the application was recorded. Deviations occur if the participant, for example, went to a wrong screen, clicked on an incorrect menu item, followed an incorrect link, or interacted incorrectly with an on-screen control. This path was compared to the optimal path. The number of steps in the observed path is divided by the number of optimal steps to provide a ratio of path deviation. It is strongly recommended that task deviations be reported. Optimal paths (i.e., procedural steps) should be recorded when constructing tasks.
Efficiency: Task Time	Each task was timed from when the administrator said "Begin" until the participant said, "Done." If he or she failed to say "Done," the time was stopped when the participant stopped performing the task. Only task times for tasks that were successfully completed were included in the average task time analysis. Average time per task was calculated for each task. Variance measures (standard deviation and standard error) were also calculated.
Satisfaction: Task Rating	Participant's subjective impression of the ease of use of the application was measured by administering both a simple post-task question as well as a post-session questionnaire. After each task, the participant was asked to rate "Overall, this task was:" on a scale of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants. Common convention is that average ratings for systems judged easy to use should be 3.3 or above. To measure participants' confidence in and likeability of the [EHRUT] overall, the testing team administered the System Usability Scale (SUS) post-test questionnaire. Questions included, "I think I would like to use this system frequently," "I thought the system was easy to use," and "I would imagine that most people would learn to use this system very quickly." See full System Usability Score questionnaire in Appendix.

RESULTS

DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above. Participants who failed to follow session and task instructions had their data excluded from the analyses.

The usability testing results for the EHRUT are detailed below. The results should be seen in light of the objectives and goals outlined in the Study Design section. The data should yield actionable results that if corrected, yield material, positive impact on user performance.

Measure 172.315(a)5	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD) (sec)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Enter/change ethnicity and sexual preference	15	100%	1.10	47.7	79.4%	0.87	4.33
Enter and Change smoking status	15	100%	0.36	32.9	82.3%	0.27	4.40
Enter cause and date of death	15	100%	1.10	34.5	86.2%	0.60	4.33

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 86%. Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80 would be considered above average.

MAJOR FINDINGS

- Major Findings –
  - Both the entering and the editing of information was simple to use.
  - The initial 'Edit' link was not obvious in the WebHIS 2.0 screen as it is in other parts of this system.
  - The 'Save' button was not obvious to an untrained user – both the position and the size/color blended in to the screen background and made it more difficult to locate.

AREAS FOR IMPROVEMENT

- Areas for Improvement
  - Modify the current 'edit' link to a more obvious user-facing button.
  - Reposition and redesign the 'save' button to make it more obvious to the user. It is not located in an area of the screen in which a Save button would typically be positioned.

EFFECTIVENESS

Overall it was found that the effectiveness of the demographic functionality – entering, modifying and saving demographic information, was very high, based on both participant satisfaction as well as the overall System Usability Scale score. The recommended modifications will make the system more efficient but likely not more effective.

EFFICIENCY

Efficiency of this program based on time:

Smoking status - the optimal time for this task was determined to be 40 seconds. The average time needed as demonstrated during the test was 32.9 seconds.

Race and Ethnicity – the optimal time for this task was determined to be 60 seconds. The average time needed as demonstrated during the test was 47.7 seconds.

Cause of Death – the optimal time for this task was determined to be 40 seconds. The average time needed as demonstrated during the test was 34.5 seconds.

Efficiency of this program based on path deviation:

Smoking status – the optimal number of clicks for this task was determined to be 6. The average number of clicks as demonstrated during the test was 6.3.

Race and Ethnicity – the optimal number of clicks for this task was determined to be 8. The average number of clicks as demonstrated during the test was 8.9.

Cause of Death – the optimal number of clicks for this task was determined to be 6. The average number of clicks as demonstrated during the test was 6.6.

SATISFACTION

Utilizing a Likert Scale of 1-5, 5 being 'very easy' to 1 being 'very difficult', the average user rating for each of the Demographic options is as follows:

Enter and Change Smoking Status	4.4
Enter and Change Ethnicity	4.3
Enter date and cause of death	4.3

Participant comments overall reflected a high satisfaction rate with the Demographic entry and editing screens, as noted in the following quotes collected from the follow up forms:

"Easy to use, but the long list of ethnicities was a bit confusing."
"Simple drop down choices in several categories."
"It took me a little longer to locate the 'save' button, but once I found it the rest of the screens were simple."

EHR Usability Test Report – Holy Name Medical Center

170.315(g)(3) - Safety Enhanced Design

Review/Update/Edit Patient Problem List (a)6

UCD Process Used: NISTIR 7741:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7741) NIST, Guide to the Processes Approach for Improving the Usability of Electronic Health Records. NIST, 29 Nov. 2010. Web. https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7741.pdf

Report Based on: NISTIR 7742:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7742) NIST, Customized Common Industry Format Template for Electronic Health Record Usability Testing. NIST, 16 Nov. 2010. Web.
https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7742.pdf

Contents

EHR USABILITY TEST REPORT – Patient Problem List 170.315(a)6

Report based on ISO/IEC 25062:2006 Common Industry Format for Usability Test Reports

Date of Usability Test: December 6, 2018
Date Usability Test was Conducted: December 6, 2018
Date Report was Prepared: April 15, 2019
Report Prepared by: Deborah Ross
Name of System Test Laboratory (STL): Holy Name Medical Center
STL Contact Person, Title and Affiliation: Deborah Ross, Clinical Analyst
STL Phone Number: 201-833-7114
STL Email Address: debross@mail.holyname.org
STL Mailing Address: 718 Teaneck Road, Teaneck, New Jersey 07666

EXECUTIVE SUMMARY

A usability test of the WebHIS Patient Problem List maintenance and review was conducted on December 6, 2018, in Teaneck, New Jersey by Holy Name Medical Center. The purpose of this test was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT).

During the usability test, fifteen (15) healthcare providers matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative tasks.

The study collected performance data on three (3) tasks typically conducted on an EHR:

- Add problems to the patient problem list
- Deactivate one of the current active problems
- Review and validate the resulting problem list

During the twenty (20) minute summative usability test, each participant was greeted by the administrator and they were instructed that they could withdraw at any time. All participants had varying levels of prior experience with the EHR. The administrator introduced the test, and instructed participants to complete a series of tasks (given one at a time) using the EHRUT. During the testing, the administrator timed the test, and along with the logger(s) recorded user performance data on paper. The administrator did not give the participant assistance in how to complete the task.

The following types of data were collected for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations
- Participant's satisfaction ratings of the system – LIKERT and System Usability Scale (SUS)

All participant data was de-identified – no correspondence could be made from the identity of the participant to the data collected. Various recommended metrics, in accordance with the examples set forth in the NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records were used to evaluate the usability of the EHRUT.

The summary data collected for **170.315(a)6** is listed in the **RESULTS** section of this document. The following is the recommended template utilized when calculating the individual results:

Measure 172.315g(3)	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Patient Problem List (a)6	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Enter patient problems	15						
Deactivate 1 problem	15						
Review resulting Problem list	15						

The overall results calculated from the **System Usability Scale (SUS)** provided to each participant scored the subjective satisfaction with the system based on performance with these tasks to be **86 %.** This scale was provided to each participant **only one time at the end of all of the tests**.

In addition to the performance data, qualitative data including both major findings and areas for improvement, based on participant anecdotal feedback will be listed in the RESULTS section as well.

INTRODUCTION

The EHRUT tested for this study was the Problem List Maintenance and Review in WebHIS 2.0. Designed to present medical information to healthcare providers in the acute care setting at Holy Name Medical Center, the EHRUT is a clinical information system used by all staff who participate in the care of our patients, on a clinical, financial or record keeping level. The usability testing attempted to represent realistic exercises and conditions.

The purpose of this study was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT). To this end, measures of effectiveness, efficiency, and user satisfaction were captured during the usability testing.

METHOD

PARTICIPANTS

A total of fifteen (15) participants were tested on the EHRUT. Participants in the test were staff nurses, nurses in clinical management, physicians and medical students. Participants work at Holy Name Medical Center in the clinical areas throughout the hospital. Participants did not participate in the requirements gathering, design or development of the EHRUT being tested. Participants were given the opportunity to have the same orientation and level of training as the actual end users would have received.

Recruited participants had a mix of backgrounds and demographic characteristics conforming to the recruitment screener. The following is a table of participants by characteristics, including demographics, professional experience, computing experience and user needs for assistive technology. Participant names were replaced with Participant IDs so that an individual's data cannot be tied back to individual identities.

	Participant ID	Gender	Age	Education	Occupation/ Role	Professional Experience (mo)	Computer Experience (mo)	Product Experience (mo)	Assistive Technology Needs
1	ID01	Female	30-39	Bachelor's Degree	Staff RN	12	180	12	-
2	ID02	Male	20-29	Bachelor's Degree	Staff RN	60	150	6	-
3	ID03	Female	60-69	Master's Degree	Nursing Director	420	360	120	-
4	ID04	Female	20-29	Bachelor's Degree	Medical Student	0	130	1	-
5	ID05	Female	40-49	Bachelor's Degree	Staff RN	84	240	84	-
6	ID06	Male	30-39	Doctorate Degree	Hospitalist	156	240	36	-
7	ID07	Female	40-49	Associate Degree	Staff RN	240	240	84	-
8	ID08	Male	20-29	Bachelor's Degree	Medical Student	0	180	1	-
9	ID09	Female	40-49	Bachelor's Degree	Staff RN	216	240	72	-
10	ID10	Female	50-59	Master's Degree	Staff RN	420	200	72	-
11	ID11	Male	40-49	Doctorate Degree	Hospitalist	182	192	36	-
12	ID12	Female	60-69	Associate Degree	Operating Room RN	456	120	60	-
13	ID13	Female	30-39	Bachelor's Degree	Staff RN	48	180	24	-
14	ID14	Female	20-29	Bachelor's Degree	Staff RN	36	180	12	-
15	ID15	Female	50-59	Associate Degree	Staff RN	324	240	96	-

Fifteen (15) participants (matching the demographics in the section on Participants) were recruited and fifteen (15) participated in the usability test. Zero (0) participants failed to show for the study.

Participants were scheduled for multiple fifteen (15) and twenty (20) minute sessions with five (5) minutes in between each session for debrief by the administrator(s) and data logger(s), and to reset systems to proper test conditions where required A spreadsheet was used to keep track of the participant schedule, and included each participant's demographic characteristics.

STUDY DESIGN

The objective of this test was to uncover areas where the application performed well – that is, effectively, efficiently, and with satisfaction – and areas where the application failed to meet the needs of the participants. The data from this test may serve as a baseline for future tests with an updated version of the same EHR and/or comparison with other EHRs provided the same tasks are used. In short, this testing serves as both a means to record or benchmark current usability, but also to identify areas where improvements must be made.

During the usability test, participants interacted with one EHR. Each participant used the system in the same location, and was provided with the same instructions. The system was evaluated for effectiveness, efficiency and satisfaction as defined by measures collected and analyzed for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

Additional information about the various measures can be found in the Usability Metrics section.

TASKS

A number of tasks were constructed that would be realistic and representative of the kinds of activities a user might do with this EHR, as follows:

- ***Enter the problems 'pressure ulcer' (399912005) and 'pneumonia' (233604007)***
- ***Inactivate the 'pneumonia' problem***
- ***Validate the list for accuracy – identify the appropriate active and inactive problems***

Tasks were selected based on their frequency of use and criticality of function. Tasks should always be constructed in light of the study objectives.

PROCEDURES

Upon arrival, participants were greeted by the test administrator. Their identity was verified and matched with their name on the participant schedule. Participants were then assigned an anonymous participant ID.

To ensure that the test ran smoothly, multiple staff members participated in this test – the administrator and several members of the hospital's simulation center, who are experienced test administrators/loggers and frequently participate in these types of exercises in our busy simulation training and testing center.

The administrator moderated the session including administering instructions and tasks. The administrator in conjunction with the data loggers monitored task times, obtained post-task rating data, and took notes on participant comments, in addition to taking notes on task success, path deviations, number and type of errors and comments.

Participants were instructed to perform the tasks (see specific instructions below):

- As quickly as possible making as few errors and deviations as possible.
- Without assistance; administrators were allowed to give immaterial guidance and clarification on tasks, but not instructions on use.
- Without using a think aloud technique.

For each task, the participants were given a written copy of the task. Task timing began once the administrator finished reading the question. The task time was stopped once the participant indicated they had successfully completed the task. Scoring is discussed below.

Participants' demographic information, task success rate, time on task, errors, deviations, verbal responses, and post-test questionnaire were recorded into a spreadsheet.

TEST LOCATION

The test facility included a quiet computer training room with tables and a desktop computer for each participant. To ensure that the environment was comfortable for users, noise levels were kept to a minimum with the ambient temperature within a normal range. All of the safety instruction and evacuation procedures were valid, in place, reviewed with and visible to the participants.

TEST ENVIRONMENT

The EHRUT would typically be used in a hospital (inpatient unit, procedural area, clinic). In this instance, the testing was conducted in the same facility on computers that were set up the same way as the clinical areas of the hospital where the application will be used. The computer being used was an **IBMi Power 7 running I5OS v7.2.** The environment in which the testing was conducted was the WebHIS 2.0 User Acceptance Testing (UAT) environment.

The participants used a mouse and keyboard when interacting with the EHRUT. The following is the configuration of the PCs used both in the test environment as well as throughout the organization in the clinical areas where desktop PCs are used:

Windows 10, dual 2.70GHz IntelCore i5 processor
8GB RAM
1600 x 900 resolution
RGB color format, SDR Color Space
6-bit
60 Hz refresh rate
23 inch monitor

The application was set up by the Holy Name Medical Center Information Systems department **Desktop Support Division**, according to the vendor's documentation, describing the system set-up and preparation.

System performance (i.e., response time) on the UAT is typically slower than that of the live environment. This was considered when calculating optimal times into the parameters for testing. Additionally, participants were instructed to not change any of the default system settings (such as control of font size).

TEST FORMS AND TOOLS

During the usability test, various documents and instruments were used, including:

- Participant demographic form
- Description of the test cases
- Objective test data capture form (for each test)
- Likert Scale for each test scenario
- System Usability Scale – one overall for each participant

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task (a)6 Enter 2 problems as SNOMED codes to Patient Problem List***
Task	***Enter the following problems for your patient: Pressure Ulcer (399912005) and Pneumonia (233604007)***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task (a)6 While still in the Problem List screen, inactivate a single problem***
Task	***Inactivate the Pneumonia problem***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task (a)6 Review problem list***
Task	***Review all lifetime problems and record active and inactive problems and verify validity based on exercises 1 and 2***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

PARTICIPANT INSTRUCTIONS

The administrator reads the following instructions aloud to the each participant:

> *Thank you for participating in this study. Your input is very important. Our* Problem List update and review test session today will last for 20 minutes with a 5 minute break *after everyone has completed the exercises. During that time you will use an instance of an electronic health record. I will ask you to complete a few tasks using this system and answer some questions. You should complete the tasks as quickly as possible making as few errors as possible. Please try to complete the tasks on your own following the instructions very closely.*
>
> *Please note that we are not testing you, we are testing the system, therefore if you have difficulty with any aspect of the tasks it likely means that something needs to be improved in the system. Please make note of it if you feel it will help us improve the usability of the system.*
>
> *We will be here in case you need general questions, but we are not able to instruct you or provide help in how to use the application. Overall, we are interested in how easy (or how difficult) this system is to use, what in it would be useful to you, and how we could improve it. Please be honest in your opinions. All of the information that you provide will be kept confidential. Your name will not be associated with any test data or comments at any time. Should you feel it necessary you are able to withdraw at any time during the testing.*

Following the procedural instructions, participants were shown the EHR and as their first task, were given 10 minutes to explore the system and ask general questions. Once this task was complete, the administrator gave the following instructions:

> *For each task, I will read the description to you and say "Begin." At that point, please perform the task and say "Done" once you believe you have successfully completed the task. I would like to request that you not talk aloud or verbalize while you are doing the tasks. I will ask you your impressions about the task once you are done.*

Participants were then given three (3) tasks to complete relating to Patient Problem List.

USABILITY METRICS

According to the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records*, EHRs should support a process that provides a high level of usability for all users. The goal is for users to interact with the system effectively, efficiently, and with an acceptable level of satisfaction. To this end, metrics for effectiveness, efficiency and user satisfaction were captured during the usability testing. The goals of the test were to assess:

1. Effectiveness of WebHIS 2.0 by measuring participant success rates and errors
2. Efficiency of WebHIS 2.0 by measuring the average task time and path deviations
3. Satisfaction with WebHIS 2.0 by measuring usability, using the System Usability Scale (SUS) overall (at the end of all the testing) and a Likert scale for ease of use for each exercise.

DATA SCORING

The following table details how tasks were scored, errors evaluated, and the time data analyzed.

Measures	Rationale and Scoring
Effectiveness: Task Success	A task was counted as a "Success" if the participant was able to achieve the correct outcome, without assistance, within the time allotted on a per task basis. The total number of successes were calculated for each task and then divided by the total number of times that task was attempted. The results are provided as a percentage. Task times were recorded for successes. Observed task times divided by the optimal time for each task is a measure of optimal efficiency. Optimal task performance time, as benchmarked by expert performance under realistic conditions, is recorded when constructing tasks. Target task times used for task times in the Moderator's Guide must be operationally defined by taking multiple measures of optimal performance and multiplying by some factor [e.g., 1.25] that allows some time buffer because the participants are presumably not trained to expert performance. Thus, if expert, optimal performance on a task was [x] seconds then allotted task time performance was [x * 1.25] seconds. This ratio should be aggregated across tasks and reported with mean and variance scores.
Effectiveness: Task Failures	If the participant abandoned the task, did not reach the correct answer or performed it incorrectly, or reached the end of the allotted time before successful completion, the task was counted as a "Failure." No task times were taken for errors. The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant. On a qualitative level, an enumeration of errors and error types should be collected.
Efficiency: Task Deviations	The participant's path (i.e., steps) through the application was recorded. Deviations occur if the participant, for example, went to a wrong screen, clicked on an incorrect menu item, followed an incorrect link, or interacted incorrectly with an on-screen control. This path was compared to the optimal path. The number of steps in the observed path is divided by the number of optimal steps to provide a ratio of path deviation. It is strongly recommended that task deviations be reported. Optimal paths (i.e., procedural steps) should be recorded when constructing tasks.
Efficiency: Task Time	Each task was timed from when the administrator said "Begin" until the participant said, "Done." If he or she failed to say "Done," the time was stopped when the participant stopped performing the task. Only task times for tasks that were successfully completed were included in the average task time analysis. Average time per task was calculated for each task. Variance measures (standard deviation and standard error) were also calculated.
Satisfaction: Task Rating	Participant's subjective impression of the ease of use of the application was measured by administering both a simple post-task question as well as a post-session questionnaire. After each task, the participant was asked to rate "Overall, this task was:" on a scale of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants. Common convention is that average ratings for systems judged easy to use should be 3.3 or above. To measure participants' confidence in and likeability of the [EHRUT] overall, the testing team administered the System Usability Scale (SUS) post-test questionnaire. Questions included, "I think I would like to use this system frequently," "I thought the system was easy to use," and "I would imagine that most people would learn to use this system very quickly." See full System Usability Score questionnaire in Appendix.

RESULTS

DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above. Participants who failed to follow session and task instructions had their data excluded from the analyses.

The usability testing results for the EHRUT are detailed below. The results should be seen in light of the objectives and goals outlined in the Study Design section. The data should yield actionable results that if corrected, yield material, positive impact on user performance.

Measure 170.315(g)3	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task Problem List (a)(6)	#	Mean (SD)	Deviations Observed/Optimal	Mean (SD)	Deviations Observed/Optimal	Mean (SD)	Mean
Enter 2 problems	15	100%	1.13	131 sec	87.33 %	1.13	4.00
Inactivate 1 problem	15	100%	1.18	25 sec	71.43%	0.53	4.13
Review/Record Problems	15	100%	1.33	25.7 sec	73.52%	1.0	4.27

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 86%. Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80 would be considered above average.

MAJOR FINDINGS

- Major Findings –
  - The problem list search returns multiple entries for a single problem. This is a function of the database used for the code search but was confusing to several users.
  - Once the user is comfortable with the tool it is very straightforward and easy to use
  - Not all of the column headings are labeled – causing some user confusion
  - The search button label was not obvious that it would allow 'adding' a new problem, especially since there is also an UpToDate search field on the same screen.
  - Comprehensive training is required to guarantee success with this application.

AREAS FOR IMPROVEMENT

- Areas for Improvement
  - Change the label on the 'search' button to 'add problem'
  - Add a heading to the 'history' column and tool tips over the plus sign to expand

- Include 'tool tips' to assist the user while working in this application
- Identify ways to filter the resulting list to minimize the 'similar' options

EFFECTIVENESS

Overall it was found that the effectiveness of the Problem List Maintenance and Review system was very high, based on both participant satisfaction as well as the overall System Usability Scale score. The recommended modifications will make the system more efficient but likely not more effective.

EFFICIENCY

The efficiency of this application based on time –

Adding a new Problem – the optimal time for this task was determined to be 150 seconds (2.5 minutes). The average time as demonstrated when evaluating the results of the testing was 131 seconds.

Deactivating an active Problem – the optimal time for this task was determined to be 35 seconds. The average time as demonstrated when evaluating the results of the testing was 25 seconds.

Reviewing the complete Problem list – the optimal time for this task was determined to be 35 seconds. The average time as demonstrated when evaluating the results of the testing was 25.7 seconds.

Efficiency of this application based on path deviations (number of clicks) -

Adding a new Problem - the optimal number of clicks for this task was determined to be 9. The average number of clicks demonstrated during the test was 11.6. This confirmed that the comments recorded by the participants did cause significant path deviations, although did not significantly impact the time for successfully completing this task.

Deactivating a Problem - the optimal number of clicks for this task was determined to be 3. The average number of clicks demonstrated during the test was 3.53.

Reviewing all problems - the optimal number of clicks for this task was determined to be 2. The average number of clicks demonstrated during this test was 2.67

SATISFACTION

Utilizing a Likert Scale of 1-5, 5 being 'very easy' to 1 being 'very difficult', the average user rating for each of the Radiology tasks is as follows:

Problem List –	Enter Problems	4.00
Problem List –	Inactivate one problem	4.13
Problem List -	Review and Validate Problem list	4.27

Participant comments overall reflected a very high satisfaction rate with the Problem List Maintenance system, as noted in the following quotes collected from the follow up forms:

“Once I understood where to click this became a very easy and useful program”.

“I’ve tried it and it is super easy”.

EHR Usability Test Report
170.315(g)(3) Safety Enhanced Design
Medication List (a)7

UCD Process Used: NISTIR 7741:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7741) NIST, Guide to the Processes Approach for Improving the Usability of Electronic Health Records. NIST, 29 Nov. 2010. Web. https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7741.pdf

Report Based on: NISTIR 7742:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7742) NIST, Customized Common Industry Format Template for Electronic Health Record Usability Testing. NIST, 16 Nov. 2010. Web.
https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7742.pdf

Contents

EHR USABILITY TEST REPORT – Medication List

Report based on ISO/IEC 25062:2006 Common Industry Format for Usability Test Reports

Date of Usability Test: July 30, 2019
Date Usability Test was Conducted: July 30, 2019
Date Report was Prepared: August 2, 2019
Report Prepared by: Deborah Ross
Name of System Test Laboratory (STL): Holy Name Medical Center
STL Contact Person, Title and Affiliation: Deborah Ross, Clinical Analyst
STL Phone Number: 201-833-7114
STL Email Address: debross@mail.holyname.org
STL Mailing Address: 718 Teaneck Road, Teaneck, New Jersey 07666

EXECUTIVE SUMMARY

A usability test of the Medication List application in WebHIS 2.0 was conducted during the week of July 29, 2019 in Teaneck, New Jersey by Holy Name Medical Center. The purpose of this test was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT).

During the usability test, ten (10) healthcare providers matching the target demographic criteria (Pharmacists) served as participants and used the EHRUT in simulated, but representative tasks.

The study collected performance data on three (3) tasks typically conducted on an EHR:

- Add the medication Metoprolol Tartrate, 50 mg, p.o. q12h, hold if heart rate < 55
- Modify the above order to 25mg, p.o. q8h, hold if heart rate <55
- Review the patient's Medication Administration Record to confirm correct display and audit trail of all actions

During the twenty (20) minute summative usability test, each participant was greeted by the administrator and they were instructed that they could withdraw at any time. All participants had varying levels of prior experience with the EHR. The administrator introduced the test, and instructed participants to complete a series of tasks (given one at a time) using the EHRUT. During the testing, the administrator timed the test, and along with the logger(s) recorded user performance data on paper. The administrator did not give the participant assistance in how to complete the task.

The following types of data were collected for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations
- Participant's satisfaction ratings of the system – LIKERT and System Usability Scale (SUS)

All participant data was de-identified – no correspondence could be made from the identity of the participant to the data collected. Various recommended metrics, in accordance with the examples set forth in the NIST Guide to the Processes Approach for Improving the Usability of Electronic Health

Records were used to evaluate the usability of the EHRUT. The summary data collected for 170.315(a)7 is listed in the **RESULTS** section of this document. The following is the recommended template utilized when calculating the individual results:

Measure 172.315g(3)	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task (a)7	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Add a new medication to the MAR (med list)	10						
Change elements of the above order	10						
Review medication list	10						

The overall results calculated from the **System Usability Scale (SUS)** provided to each participant scored the subjective satisfaction with the system based on performance with these tasks to be **76%.** This scale was provided to each participant **only one time at the end of all of the tests**.

In addition to the performance data, qualitative data including both major findings and areas for improvement, based on participant anecdotal feedback will be listed in the RESULTS section as well.

INTRODUCTION

The EHRUT tested for this study utilized the Medication Profiling system in WebHIS 2.0, and RUMBA. Designed to present medical information to healthcare providers in the acute care setting at Holy Name Medical Center, the EHRUT is a clinical information system used by all staff who participate in the care of our patients, on a clinical, financial or record keeping level. The usability testing attempted to represent realistic exercises and conditions.

The purpose of this study was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT). To this end, measures of effectiveness, efficiency, and user satisfaction were captured during the usability testing.

METHOD

PARTICIPANTS

A total of ten (10) participants were tested on the EHRUT. Participants in the test were Staff Pharmacists, Clinical Pharmacy Specialists and Pharmacy Informaticists. Participants did not participate in the requirements gathering, design or development of the EHRUT being tested. Participants were

given the opportunity to have the same orientation and level of training as the actual end users would have received.

Recruited participants had a mix of backgrounds and demographic characteristics conforming to the recruitment screener. The following is a table of participants by characteristics, including demographics, professional experience, computing experience and user needs for assistive technology. Participant names were replaced with Participant IDs so that an individual's data cannot be tied back to individual identities.

	Participant ID	Gender	Age	Education	Occupation/ Role	Professional Experience (mo)	Computer Experience (mo)	Product Experience (mo)	Assistive Technology Needs
16	ID01	Female	30-39	Doctorate Degree	Clinical Pharmacy Specialist-Informatics	84	252	36	-
17	ID02	Female	40-49	Doctorate Degree	Staff Pharmacist	144	300	132	-
18	ID03	Male	30-39	Doctorate Degree	Clin.Pharmacy Specialist	84	300	120	-
19	ID04	Male	20-29	Doctorate Degree	Clin Pharmacy Specialist Oncology	24	240	24	-
20	ID05	Female	20-29	Doctorate Degree	Staff Pharmacist	12	168	12	-
21	ID06	Female	60-69	Bachelor's Degree	Staff Pharmacist	480	360	120	-
22	ID07	Female	30-39	Bachelor's Degree	Staff Pharmacist	54	336	30	-
23	ID08	Male	20-29	Doctorate Degree	Staff Pharmacist Oncology	30	240	30	-
24	ID09	Male	30-39	Doctorate Degree	Pharmacy Operations Manager	108	324	192	-
25	ID10	Male	50-59	Bachelor's Degree	Manager Outpatient Pharmacy Services	396	360	192	-

Ten (10) participants (matching the demographics in the section on Participants) were recruited and ten (10) participated in the usability test. Zero (0) participants failed to show for the study.

Participants were scheduled for a twenty (20) minute session to complete this test.

STUDY DESIGN

The objective of this test was to uncover areas where the application performed well – that is, effectively, efficiently, and with satisfaction – and areas where the application failed to meet the needs of the participants. The data from this test may serve as a baseline for future tests with an updated version of the same EHR and/or comparison with other EHRs provided the same tasks are used. In short, this testing serves as both a means to record or benchmark current usability, but also to identify areas where improvements must be made.

During the usability test, participants interacted with one EHR. Each participant used the system in the same location, and was provided with the same instructions. The system was evaluated for effectiveness, efficiency and satisfaction as defined by measures collected and analyzed for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

Additional information about the various measures can be found in the section on Usability Metrics.

TASKS

A number of tasks were constructed that would be realistic and representative of the kinds of activities a user might do with this EHR, as follows:

- ***Profile an order for Metoprolol Tartrate 50 mg every 12 hours, hold for a heart rate less than 55.***
- ***Modify the above order to Metoprolol Tartrate 25 mg every 8 hours, hold for a heart rate less than 55.***
- ***Review the patient's medication list and identify the profiled medication and history of modifications***

Tasks were selected based on their frequency of use, criticality of function, and those that may be most troublesome for users. Tasks should always be constructed in light of the study objectives.

PROCEDURES

Upon arrival, participants were greeted by the test administrator. Their identity was verified and matched with their name on the participant schedule. Participants were then assigned an anonymous participant ID.

To ensure that the test ran smoothly, multiple staff members participated in this test – the administrator and several members of the hospital's simulation center, who are experienced test administrators/loggers and frequently participate in these types of exercises in our busy simulation training and testing center.

The administrator moderated the session including administering instructions and tasks. The administrator in conjunction with the data loggers monitored task times, obtained post-task rating data, and took notes on participant comments, in addition to taking notes on task success, path deviations, number and type of errors and comments.

Participants were instructed to perform the tasks (see specific instructions below):

- As quickly as possible making as few errors and deviations as possible.
- Without assistance; administrators were allowed to give immaterial guidance and clarification on tasks, but not instructions on use.
- Without using a think aloud technique.

For each task, the participants were given a written copy of the task. Task timing began once the administrator finished reading the question. The task time was stopped once the participant indicated they had successfully completed the task. Scoring is discussed below.

Participants' demographic information, task success rate, time on task, errors, deviations, verbal responses, and post-test questionnaire were recorded into a spreadsheet.

TEST LOCATION

The test facility included a quiet computer training room with tables and a desktop computer for each participant. To ensure that the environment was comfortable for users, noise levels were kept to a minimum with the ambient temperature within a normal range. All of the safety instruction and evacuation procedures were valid, in place, reviewed with and visible to the participants.

TEST ENVIRONMENT

The EHRUT would typically be used in a hospital (inpatient unit, procedural area, clinic). In this instance, the testing was conducted in the same facility on computers that were set up the same way as the clinical areas of the hospital where the application will be used. The computer being used was an **IBMi Power 7 running I5OS v7.2.** The environment in which the testing was conducted was the WebHIS 2.0 User Acceptance Testing (UAT) environment.

The participants used a mouse and keyboard when interacting with the EHRUT. The following is the configuration of the PCs used both in the test environment as well as throughout the organization in the clinical areas where desktop PCs are used:

Windows 10, dual 2.70GHz IntelCore i5 processor
8GB RAM
1600 x 900 resolution
RGB color format, SDR Color Space
6-bit
60 Hz refresh rate
23 inch monitor

The application was set up by the Holy Name Medical Center Information Systems department '**Desktop Support Division'**, according to the vendor's documentation, describing the system set-up and preparation.

System performance (i.e., response time) on the UAT is typically slower than that of the live environment. This was considered when calculating optimal times into the parameters for testing. Additionally, participants were instructed to not change any of the default system settings (such as control of font size).

TEST FORMS AND TOOLS

During the usability test, various documents and instruments were used, including:

- Participant demographic form
- Description of the test cases
- Objective test data capture form (for each test)
- Likert Scale for each test scenario
- System Usability Scale – one overall for each participant

Sample Objective Test Data Forms:

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 7(a) Medication List***
Task	***Profile the medication Metoprolol Tartrate 50 mg PO q12 hours; hold for heart rate less than 55.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 7(a) Medication List***
Task	***Modify the above order to Metoprolol Tartrate 25mg every 8 hours; hold for heart rate less than 55.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	WebHIS 2.0 Patient List ***Task 7(a) Medication List***
Task	***View Medication list and identify newly added/modified medications***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

PARTICIPANT INSTRUCTIONS

The administrator reads the following instructions aloud to the each participant:

> *Thank you for participating in this study. Your input is very important. Our Medication List test session today will last for 20 minutes. During that time you will use an instance of an electronic health record. I will ask you to complete a few tasks using this system and answer some questions. You should complete the tasks as quickly as possible making as few errors as possible. Please try to complete the tasks on your own following the instructions very closely.*
>
> *Please note that we are not testing you, we are testing the system, therefore if you have difficulty with any aspect of the tasks it likely means that something needs to be improved in the system. Please make note of it if you feel it will help us improve the usability of the system.*
>
> *We will be here in case you need general questions, but we are not able to instruct you or provide help in how to use the application. Overall, we are interested in how easy (or how difficult) this system is to use, what in it would be useful to you, and how we could improve it. Please be honest in your opinions. All of the information that you provide will be kept confidential. Your name will not be associated with any test data or comments at any time. Should you feel it necessary you are able to withdraw at any time during the testing.*

Following the procedural instructions, participants were shown the EHR and as their first task, were given 10 minutes to explore the system and ask general questions. Once this task was complete, the administrator gave the following instructions:

> *For each task, I will read the description to you and say "Begin." At that point, please perform the task and say "Done" once you believe you have successfully completed the task. I would like to request that you not talk aloud or verbalize while you are doing the tasks. I will ask you your impressions about the task once you are done.*

Participants were then given three (3) tasks to complete relating to creating, modifying and reviewing a patient's Medication List.

USABILITY METRICS

According to the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records*, EHRs should support a process that provides a high level of usability for all users. The goal is for users to interact with the system effectively, efficiently, and with an acceptable level of satisfaction. To this end, metrics for effectiveness, efficiency and user satisfaction were captured during the usability testing. The goals of the test were to assess:

1. Effectiveness of WebHIS 2.0 by measuring participant success rates and errors
2. Efficiency of WebHIS 2.0 by measuring the average task time and path deviations
3. Satisfaction with WebHIS 2.0 by measuring usability, using the System Usability Scale (SUS) overall (at the end of all the testing) and a Likert scale for ease of use for each exercise.

DATA SCORING

The following table details how tasks were scored, errors evaluated, and the time data analyzed.

Measures	Rationale and Scoring
Effectiveness: Task Success	A task was counted as a "Success" if the participant was able to achieve the correct outcome, without assistance, within the time allotted on a per task basis. The total number of successes were calculated for each task and then divided by the total number of times that task was attempted. The results are provided as a percentage. Task times were recorded for successes. Observed task times divided by the optimal time for each task is a measure of optimal efficiency. Optimal task performance time, as benchmarked by expert performance under realistic conditions, is recorded when constructing tasks. Target task times used for task times in the Moderator's Guide must be operationally defined by taking multiple measures of optimal performance and multiplying by some factor [e.g., 1.25] that allows some time buffer because the participants are presumably not trained to expert performance. Thus, if expert, optimal performance on a task was [x] seconds then allotted task time performance was [x * 1.25] seconds. This ratio should be aggregated across tasks and reported with mean and variance scores.
Effectiveness: Task Failures	If the participant abandoned the task, did not reach the correct answer or performed it incorrectly, or reached the end of the allotted time before successful completion, the task was counted as a "Failure." No task times were taken for errors. The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant. On a qualitative level, an enumeration of errors and error types should be collected.
Efficiency: Task Deviations	The participant's path (i.e., steps) through the application was recorded. Deviations occur if the participant, for example, went to a wrong screen, clicked on an incorrect menu item, followed an incorrect link, or interacted incorrectly with an on-screen control. This path was compared to the optimal path. The number of steps in the observed path is divided by the number of optimal steps to provide a ratio of path deviation. It is strongly recommended that task deviations be reported. Optimal paths (i.e., procedural steps) should be recorded when constructing tasks.
Efficiency: Task Time	Each task was timed from when the administrator said "Begin" until the participant said, "Done." If he or she failed to say "Done," the time was stopped when the participant stopped performing the task. Only task times for tasks that were successfully completed were included in the average task time analysis. Average time per task was calculated for each task. Variance measures (standard deviation and standard error) were also calculated.
Satisfaction: Task Rating	Participant's subjective impression of the ease of use of the application was measured by administering both a simple post-task question as well as a post-session questionnaire. After each task, the participant was asked to rate "Overall, this task was:" on a scale of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants. Common convention is that average ratings for systems judged easy to use should be 3.3 or above. To measure participants' confidence in and likeability of the [EHRUT] overall, the testing team administered the System Usability Scale (SUS) post-test questionnaire. Questions included, "I think I would like to use this system frequently," "I thought the system was easy to use," and "I would imagine that most people would learn to use this system very quickly." See full System Usability Score questionnaire in Appendix.

RESULTS

DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above. Participants who failed to follow session and task instructions had their data excluded from the analyses.

The usability testing results for the EHRUT are detailed below. The results should be seen in light of the objectives and goals outlined in the Study Design section. The data should yield actionable results that if corrected, yield material, positive impact on user performance.

Measure 172.315(g)3	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task (a)7	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Enter new medication into medication list		100%	1.11	48.20	64.27%	1.10	4.20
Modify the above medication		100%	1.23	33.00	66.00%	2.30	3.90
Review the medication list		100%	1.00	21.70	43.40%	0	4.70

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 86%. Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80 would be considered above average.

MAJOR FINDINGS

- Major Findings –
  - Difficult to work through the screens without prior training
  - Too many redundant clicks
  - Some users stated that it was easy to figure out but involved too many clicks
  - Adding parameters for orders was time consuming
  - Function keys should be defined on the screen
  - WebMAR is not user friendly – no legend, etc. apparent to users
  - No functionality to edit a medication – must either copy or enter an entirely new order
  - Copy function does not do interaction checking

AREAS FOR IMPROVEMENT

- Areas for Improvement
  - Add legend for function keys to RXO program
  - Add full function order modification program – including interaction checking
  - Streamline functionality to decrease number of clicks

EFFECTIVENESS

Overall it was determined that the effectiveness of the Medication List program (building and modifying the medication list) was high in the respect that the lists were built accurately but as evident from the participants' comments the process was time consuming and redundant.

EFFICIENCY

The efficiency of this application based on time –

Enter a new medication order – the optimal time in which to complete this task was calculated to be 75 seconds. The average time required to complete this task as demonstrated during the test was 48.2 seconds.

Modify a medication order - the optimal time for this task was determined to be 50 seconds. The average time required to complete this test was 33 seconds.

Review the medication list - the optimal time for this task was determined to be 50 seconds. The average time required to complete this test was 21.70 seconds.

Efficiency of this application based on path deviations (number of clicks) -

Enter a new medication order - the optimal number of clicks for this task was determined to be 10. The average number of clicks required to perform this function as demonstrated during the test was 11.1.

Modify a medication order – the optimal number of clicks for this task was determined to be 10. The average number of clicks required to perform this function as demonstrated during this test was 12.3.

Review the medication list - the optimal number of clicks for this task was determined to be 6. The average number of clicks required to perform this function as demonstrated during this test was 6.

SATISFACTION

Utilizing a Likert Scale of 1-5, 5 being 'very easy' to 1 being 'very difficult', the average user rating for each of the Drug Alert options is as follows:

Adding (profiling) a new Medication	4.20
Modifying a Medication	3.90
Reviewing Medication	4.70

Participant comments reflected a mixed satisfaction rate with the Medication profiling and modification program, as noted in the following quotes collected from the follow up forms:

"The system is easy to figure out but it requires many redundant 'clicks'"

“There is no functionality that allows order modification - it is too time consuming and requires changes. We must enter a new order for most modifications.”

“The tasks assigned were clear and simple.”

“We have to work between systems in order to view the final medication list. This is not efficient and we aren’t familiar with that system.”

“Please add more prompts to the screens in these programs.”

EHR Usability Test Report – 170.315(g)(3) Safety Enhanced Design Medication Allergy List (a)8

UCD Process Used: NISTIR 7741:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7741) NIST, Guide to the Processes Approach for Improving the Usability of Electronic Health Records. NIST, 29 Nov. 2010. Web. https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7741.pdf

Report Based on: NISTIR 7742:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7742) NIST, Customized Common Industry Format Template for Electronic Health Record Usability Testing. NIST, 16 Nov. 2010. Web.
https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7742.pdf

Contents

EHR USABILITY TEST REPORT – Medication Allergy List(a)8

Report based on ISO/IEC 25062:2006 Common Industry Format for Usability Test Reports

Date of Usability Test: December 6, 2018
Date Usability Test was Conducted: December 6, 2018
Date Report was Prepared: April 15, 2019
Report Prepared by: Deborah Ross
Name of System Test Laboratory (STL): Holy Name Medical Center
STL Contact Person, Title and Affiliation: Deborah Ross, Clinical Analyst
STL Phone Number: 201-833-7114
STL Email Address: debross@mail.holyname.org
STL Mailing Address: 718 Teaneck Road, Teaneck, New Jersey 07666

EXECUTIVE SUMMARY

A usability test of the Medication Allergy List program in WebHIS 2.0 was conducted on December 6, 2018, in Teaneck, New Jersey by Holy Name Medical Center. The purpose of this test was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT).

During the usability test, fifteen (15) healthcare providers matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative tasks.

The study collected performance data and user reactions to multiple tasks related to documenting and retrieving information regarding a client's medication allergy list:

- Record a Medication Allergy
- Change a Medication Allergy
- Access and display the current Medication Allergy List

During the twenty (20) minute summative usability test, each participant was greeted by the administrator and they were instructed that they could withdraw at any time. All participants had varying levels of prior experience with the EHR. The administrator introduced the test, and instructed participants to complete a series of tasks (given one at a time) using the EHRUT. During the testing, the administrator timed the test, and along with the logger(s) recorded user performance data on paper. The administrator did not give the participant assistance in how to complete the task.

The following types of data were collected for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations
- Participant's satisfaction ratings of the system – LIKERT and System Usability Scale (SUS)

All participant data was de-identified – no correspondence could be made from the identity of the participant to the data collected. Various recommended metrics, in accordance with the examples set forth in the NIST Guide to the Processes Approach for Improving the Usability of Electronic Health

Records were used to evaluate the usability of the EHRUT. The summary data collected for *170.315(a)8* is listed in the **RESULTS** section of this document. The following is the recommended template utilized when calculating the individual results:

Measure 170.315(g)3	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
(a)8 Task	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Record Medication Allergy	15						
Change Medication Allergy	15						
Access Medication Allergy List	15						

The overall results calculated from the **System Usability Scale (SUS)** provided to each participant scored the subjective satisfaction with the system based on performance with these tasks to be **86 %.** This scale was provided to each participant **only one time at the end of all of the tests**.

In addition to the performance data, qualitative data including both major findings and areas for improvement, based on participant anecdotal feedback will be listed in the RESULTS section as well.

INTRODUCTION

The EHRUT tested for this study were the various tasks associated with recording, updating and reviewing a patient's Medication Allergies in WebHIS 2.0.

Designed to present medical information to healthcare providers in the acute care setting at Holy Name Medical Center, the EHRUT is a clinical information system used by all staff who participate in the care of our patients, on a clinical, financial or record keeping level. The usability testing attempted to represent realistic exercises and conditions.

The purpose of this study was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT). To this end, measures of effectiveness, efficiency, and user satisfaction were captured during the usability testing.

METHOD

PARTICIPANTS

A total of fifteen (15) participants were tested on the EHRUT. Participants in the test were staff nurses, nurses in clinical management, physicians and medical students. Participants work at Holy Name Medical Center in the clinical areas throughout the hospital. Participants did not participate in the requirements gathering, design or development of the EHRUT being tested. Participants were given the

opportunity to have the same orientation and level of training as the actual end users would have received.

Recruited participants had a mix of backgrounds and demographic characteristics conforming to the recruitment screener. The following is a table of participants by characteristics, including demographics, professional experience, computing experience and user needs for assistive technology. Participant names were replaced with Participant IDs so that an individual's data cannot be tied back to individual identities.

	Participant ID	Gender	Age	Education	Occupation/ Role	Professional Experience (mo)	Computer Experience (mo)	Product Experience (mo)	Assistive Technology Needs
1	ID01	Female	30-39	Bachelor's Degree	Staff RN	12	180	12	-
2	ID02	Male	20-29	Bachelor's Degree	Staff RN	60	150	6	-
3	ID03	Female	60-69	Master's Degree	Nursing Director	420	360	120	-
4	ID04	Female	20-29	Bachelor's Degree	Medical Student	0	130	1	-
5	ID05	Female	40-49	Bachelor's Degree	Staff RN	84	240	84	-
6	ID06	Male	30-39	Doctorate Degree	Hospitalist	156	240	36	-
7	ID07	Female	40-49	Associate Degree	Staff RN	240	240	84	-
8	ID08	Male	20-29	Bachelor's Degree	Medical Student	0	180	1	-
9	ID09	Female	40-49	Bachelor's Degree	Staff RN	216	240	72	-
10	ID10	Female	50-59	Master's Degree	Staff RN	420	200	72	-
11	ID11	Male	40-49	Doctorate Degree	Hospitalist	182	192	36	-
12	ID12	Female	60-69	Associate Degree	Operating Room RN	456	120	60	-
13	ID13	Female	30-39	Bachelor's Degree	Staff RN	48	180	24	-
14	ID14	Female	20-29	Bachelor's Degree	Staff RN	36	180	12	-
15	ID15	Female	50-59	Associate Degree	Staff RN	324	240	96	-

Fifteen (15) participants (matching the demographics in the section on Participants) were recruited and fifteen (15) participated in the usability test. Zero (0) participants failed to show for the study.

Participants were scheduled for multiple fifteen (15) and twenty (20) minute sessions with five (5) minutes in between each session for debrief by the administrator(s) and data logger(s), and to reset

systems to proper test conditions where required A spreadsheet was used to keep track of the participant schedule, and included each participant's demographic characteristics.

STUDY DESIGN

The objective of this test was to uncover areas where the application performed well – that is, effectively, efficiently, and with satisfaction – and areas where the application failed to meet the needs of the participants. The data from this test may serve as a baseline for future tests with an updated version of the same EHR and/or comparison with other EHRs provided the same tasks are used. In short, this testing serves as both a means to record or benchmark current usability, but also to identify areas where improvements must be made.

During the usability test, participants interacted with one EHR. Each participant used the system in the same location, and was provided with the same instructions. The system was evaluated for effectiveness, efficiency and satisfaction as defined by measures collected and analyzed for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

Additional information about the various measures can be found in the section on Usability Metrics.

TASKS

A number of tasks were constructed that would be realistic and representative of the kinds of activities a user might do with this EHR, including but not limited to:

- ***Documenting a newly identified Medication Allergy***
- ***Updating the status of a currently documented Medication Allergy***
- ***Reviewing the complete list of active and inactive patient Medication Allergies***

These are realistic scenarios that are used during procedures and surgeries on a daily basis, whether implanting or removing a said device, implant or prosthetic.

PROCEDURES

Upon arrival, participants were greeted by the test administrator. Their identity was verified and matched with their name on the participant schedule. Participants were then assigned an anonymous participant ID.

To ensure that the test ran smoothly, multiple staff members participated in this test – the administrator and several members of the hospital's simulation center, who are experienced test administrators/loggers and frequently participate in these types of exercises in our busy simulation training and testing center.

The administrator moderated the session including administering instructions and tasks. The administrator in conjunction with the data loggers monitored task times, obtained post-task rating data, and took notes on participant comments, in addition to taking notes on task success, path deviations, number and type of errors and comments.

Participants were instructed to perform the tasks (see specific instructions below):

- As quickly as possible making as few errors and deviations as possible.
- Without assistance; administrators were allowed to give immaterial guidance and clarification on tasks, but not instructions on use.
- Without using a think aloud technique.

For each task, the participants were given a written copy of the task. Task timing began once the administrator finished reading the question. The task time was stopped once the participant indicated they had successfully completed the task. Scoring is discussed below.

Participants' demographic information, task success rate, time on task, errors, deviations, verbal responses, and post-test questionnaire were recorded into a spreadsheet.

TEST LOCATION

The test facility included a quiet computer training room with tables and a desktop computer for each participant. To ensure that the environment was comfortable for users, noise levels were kept to a minimum with the ambient temperature within a normal range. All of the safety instruction and evacuation procedures were valid, in place, reviewed with and visible to the participants.

TEST ENVIRONMENT

The EHRUT would typically be used in a hospital (inpatient unit, procedural area, clinic). In this instance, the testing was conducted in the same facility on computers that were set up the same way as the clinical areas of the hospital where the application will be used. The computer being used was an **IBMi Power 7 running I5OS v7.2.** The environment in which the testing was conducted was the WebHIS 2.0 User Acceptance Testing (UAT) environment.

The participants used a mouse, keyboard and barcode scanner when interacting with the EHRUT. The following is the configuration of the PCs used both in the test environment as well as throughout the organization in the clinical areas where desktop PCs are used:

Windows 10, dual 2.70GHz IntelCore i5 processor
8GB RAM
1600 x 900 resolution
RGB color format, SDR Color Space
6-bit
60 Hz refresh rate
23 inch monitor
Ds8100-HC Series Handheld Imager (for barcode scanning)

The application was set up by the Holy Name Medical Center Information Systems department **Desktop Support Division**, according to the vendor's documentation, describing the system set-up and preparation.

System performance (i.e., response time) on the UAT is typically slower than that of the live environment. This was considered when calculating optimal times into the parameters for testing. Additionally, participants were instructed to not change any of the default system settings (such as control of font size).

TEST FORMS AND TOOLS

During the usability test, various documents and instruments were used, including:

- Participant demographic form
- Description of the test cases
- Objective test data capture form (for each test)
- Likert Scale for each test scenario
- System Usability Scale – one overall for each participant

Sample Objective Test Data Forms:

Task Performance Evaluation	**WebHIS 2.0** **Task 175.310(a)8 Medication Allergy List**
Task	***Record the allergy 'penicillin' for your patient with reaction 'shock'***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	**WebHIS 2.0** **Task 170.315(a)8 Medication Allergy List**
Task	***Inactivate the penicillin allergy for your patient***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	***WebHIS 2.0*** ***Task 170.315(a)8 – Patient Allergy List***
Task	***Review the list of active and inactive medication allergies for your patient. Note the allergy that you added and updated.***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

PARTICIPANT INSTRUCTIONS

The administrator reads the following instructions aloud to the each participant:

> *Thank you for participating in this study. Your input is very important. The Medication Allergy test session today will last for twenty minutes with a five minute break after everyone has completed the exercises. During that time you will use an instance of an electronic health record. I will ask you to complete a few tasks using this system and answer some questions. You should complete the tasks as quickly as possible making as few errors as possible. Please try to complete the tasks on your own following the instructions very closely.*
>
> *Please note that we are not testing you, we are testing the system, therefore if you have difficulty with any aspect of the tasks it likely means that something needs to be improved in the system. Please make note of it if you feel it will help us improve the usability of the system.*
>
> *We will be here in case you have general questions, but we are not able to instruct you or provide help in how to use the application. Overall, we are interested in how easy (or how difficult) this system is to use, what in it would be useful to you, and how we could improve it. Please be honest in your opinions. All of the information that you provide will be kept confidential. Your name will not be associated with any test data or comments at any time. Should you feel it necessary you are able to withdraw at any time during the testing.*

Following the procedural instructions, participants were shown the EHR and as their first task, were given 10 minutes to explore the system and ask general questions. Once this task was complete, the administrator gave the following instructions:

> *For each task, I will read the description to you and say "Begin." At that point, please perform the task and say "Done" once you believe you have successfully completed the task. I would like to request that you not talk aloud or verbalize while you are doing the tasks. I will ask you your impressions about the task once you are done.*

Participants were then given three (3) tasks within the Medication Allergy program to complete and evaluate.

USABILITY METRICS

According to the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records*, EHRs should support a process that provides a high level of usability for all users. The goal is for users to interact with the system effectively, efficiently, and with an acceptable level of satisfaction. To this end, metrics for effectiveness, efficiency and user satisfaction were captured during the usability testing. The goals of the test were to assess:

1. Effectiveness of WebHIS 2.0 by measuring participant success rates and errors
2. Efficiency of WebHIS 2.0 by measuring the average task time and path deviations
3. Satisfaction with WebHIS 2.0 by measuring usability, using the System Usability Scale (SUS) overall (at the end of all the testing) and a Likert scale for ease of use for each exercise.

DATA SCORING

The following table details how tasks were scored, errors evaluated, and the time data analyzed.

Measures	Rationale and Scoring
Effectiveness: Task Success	A task was counted as a "Success" if the participant was able to achieve the correct outcome, without assistance, within the time allotted on a per task basis. The total number of successes were calculated for each task and then divided by the total number of times that task was attempted. The results are provided as a percentage. Task times were recorded for successes. Observed task times divided by the optimal time for each task is a measure of optimal efficiency. Optimal task performance time, as benchmarked by expert performance under realistic conditions, is recorded when constructing tasks. Target task times used for task times in the Moderator's Guide must be operationally defined by taking multiple measures of optimal performance and multiplying by some factor [e.g., 1.25] that allows some time buffer because the participants are presumably not trained to expert performance. Thus, if expert, optimal performance on a task was [x] seconds then allotted task time performance was [x * 1.25] seconds. This ratio should be aggregated across tasks and reported with mean and variance scores.
Effectiveness: Task Failures	If the participant abandoned the task, did not reach the correct answer or performed it incorrectly, or reached the end of the allotted time before successful completion, the task was counted as a "Failure." No task times were taken for errors. The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant. On a qualitative level, an enumeration of errors and error types should be collected.
Efficiency: Task Deviations	The participant's path (i.e., steps) through the application was recorded. Deviations occur if the participant, for example, went to a wrong screen, clicked on an incorrect menu item, followed an incorrect link, or interacted incorrectly with an on-screen control. This path was compared to the optimal path. The number of steps in the observed path is divided by the number of optimal steps to provide a ratio of path deviation. It is strongly recommended that task deviations be reported. Optimal paths (i.e., procedural steps) should be recorded when constructing tasks.
Efficiency: Task Time	Each task was timed from when the administrator said "Begin" until the participant said, "Done." If he or she failed to say "Done," the time was stopped when the participant stopped performing the task. Only task times for tasks that were successfully completed were included in the average task time analysis. Average time per task was calculated for each task. Variance measures (standard deviation and standard error) were also calculated.
Satisfaction: Task Rating	Participant's subjective impression of the ease of use of the application was measured by administering both a simple post-task question as well as a post-session questionnaire. After each task, the participant was asked to rate "Overall, this task was:" on a scale of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants. Common convention is that average ratings for systems judged easy to use should be 3.3 or above. To measure participants' confidence in and likeability of the [EHRUT] overall, the testing team administered the System Usability Scale (SUS) post-test questionnaire. Questions included, "I think I would like to use this system frequently," "I thought the system was easy to use," and "I would imagine that most people would learn to use this system very quickly." See full System Usability Score questionnaire in Appendix.

RESULTS

DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above. Participants who failed to follow session and task instructions had their data excluded from the analyses.

The usability testing results for the EHRUT are detailed below. The results should be seen in light of the objectives and goals outlined in the Study Design section. The data should yield actionable results that if corrected, yield material, positive impact on user performance.

Measure 170.315(g)3	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
(a)14 Task	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Record Medication Allergy	15	100%	1.16	28.1 sec	70.2%	1.4	4.13
Change the Medication Allergy Status	15	100%	1.25	20.6 sec	68.7%	1.27	4.20
Review the complete Medication Allergy List	15	100%	1.30	19.4 sec	77.6%	1.20	4.20

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 86%. Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80 would be considered above average.

MAJOR FINDINGS

- Major Findings –
    - The 'maintain allergies' button was not descriptive to the tasks it represents
    - The medication search took a while to respond removal, which is a mandatory
    - The medication search returns too many options for each substance – it should only return one response per medication name.
    - When viewing the initial list there is no way to view the audit trail and history – the user must go to the 'maintain allergies' view.

AREAS FOR IMPROVEMENT

- Areas for Improvement

    - Change the 'Maintain Allergies' label to read 'Add/Edit Allergies'
    - Make that button a different color to draw the user to it
    - Include a toggle for 'History' (the audit trail) on the initial view and all subsequent views
    - Reduce the medication list to one instance for each substance – additional forms are redundant in this application. This will speed up the search significantly.

EFFECTIVENESS

Overall it was found that the effectiveness of the Medication Allergy program was high. Although users did make valuable observations and recommendations regarding the User Interface, the application itself captures and displays current and inactive medication allergies and alerts patients to said allergies in applications where the information is needed for patient safety and clinical decision support, such as in CPOE. With some minor cosmetic changes as recommended above this application will continue to support patient safety and be easy to use.

EFFICIENCY

The efficiency of this application based on time –

Enter new medication allergy – the optimal time for this task was determined to be 40 seconds. The average time as demonstrated during the test was 28.1 seconds.

Modify the medication allergy – the optimal time for this task was determined to be 30 seconds. The average time as demonstrated during the test was 20.6 seconds.

Review the list of current and deactivated medication allergies – the optimal time for this task was determined to be 25 seconds. The average time as demonstrated during the test was 19.4 seconds.

Efficiency of this application based on path deviations (number of clicks) -

Enter new medication allergy – the optimal number of clicks for this task was determined to be 9. The average number of clicks demonstrated during the test was 10.4.

Modify the medication allergy – the optimal number of clicks for this task was determined to be 5. The average number of clicks demonstrated during this test was 6.27.

Review the list of current and deactivated medication allergies – the optimal number of clicks for this task was determined to be 4. The average number of clicks demonstrated during this test was 5.2.

SATISFACTION

Utilizing a Likert Scale of 1-5, 5 being 'very easy' to 1 being 'very difficult', the average user rating for each of the Clinical Decision Support Rules was as follows:

Adding Medication Allergy	**4.13**
Change the status of the Allergy	**4.20**
Review Patient Allergy List	**4.20**

PARTICIPANT COMMENTS

Participant comments overall reflected a high satisfaction rate with the Medication Allergy application, as noted in the following quotes collected from the follow up forms:

"Easy and straight forward"
"Easy and intuitive to use"
"Easy to navigate – will be easier if some of the buttons are renamed"
"If the medication list is shorter this will be faster to use, but overall it was easy and intuitive"

EHR Usability Test Report
170.315(g)(3) Safety Enhanced Design
Implantable Devices (a)14

UCD Process Used: NISTIR 7741:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7741) NIST, Guide to the Processes Approach for Improving the Usability of Electronic Health Records. NIST, 29 Nov. 2010. Web. https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7741.pdf

Report Based on: NISTIR 7742:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7742) NIST, Customized Common Industry Format Template for Electronic Health Record Usability Testing. NIST, 16 Nov. 2010. Web.
https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7742.pdf

Contents

EHR USABILITY TEST REPORT – Implantable Devices 170.315(a)14

Report based on ISO/IEC 25062:2006 Common Industry Format for Usability Test Reports

Date of Usability Test: December 6, 2018
Date Usability Test was Conducted: December 6, 2018
Date Report was Prepared: April 15, 2019
Report Prepared by: Deborah Ross
Name of System Test Laboratory (STL): Holy Name Medical Center
STL Contact Person, Title and Affiliation: Deborah Ross, Clinical Analyst
STL Phone Number: 201-833-7114
STL Email Address: debross@mail.holyname.org
STL Mailing Address: 718 Teaneck Road, Teaneck, New Jersey 07666

EXECUTIVE SUMMARY

A usability test of the Implantable Device tracking application in WebHIS 2.0 was conducted on December 6, 2018, in Teaneck, New Jersey by Holy Name Medical Center. The purpose of this test was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT).

During the usability test, fifteen (15) healthcare providers matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative tasks.

The study collected performance data and user reactions to multiple tasks related to documenting and retrieving information regarding Implantable Devices as follows:

- Record a Unique Device Identifier (UDI) and status for a new implant for your patient
- Change the status of the implant
- Access and identify UDIs, device description, UDI identifiers and status

During the twenty (20) minute summative usability test, each participant was greeted by the administrator and they were instructed that they could withdraw from the test at any time. All participants had varying levels of prior experience with the EHR. The administrator introduced the test, and instructed participants to complete a series of tasks (given one at a time) using the EHRUT. During the testing, the administrator timed the test, and along with the logger(s) recorded user performance data on paper. The administrator did not give the participant assistance in how to complete the task.

The following types of data were collected for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations
- Participant's satisfaction ratings of the system – LIKERT and System Usability Scale (SUS)

All participant data was de-identified – no correspondence could be made from the identity of the participant to the data collected. Various recommended metrics, in accordance with the examples set forth in the NIST Guide to the Processes Approach for Improving the Usability of Electronic Health

Records were used to evaluate the usability of the EHRUT. The summary data collected for *170.315(a)14* is listed in the **RESULTS** section of this document. The following is the recommended template utilized when calculating the individual results:

Measure 170.315(g)3 (a)14	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
Task	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Record UDI and status	15						
Change status of implant	15						
Access UDI attributes	15						

The overall results calculated from the **System Usability Scale (SUS)** provided to each participant scored the subjective satisfaction with the system based on performance with these tasks to be **86 %.** This scale was provided to each participant **only one time at the end of all of the tests**.

In addition to the performance data, qualitative data including both major findings and areas for improvement, based on participant anecdotal feedback will be listed in the RESULTS section as well.

INTRODUCTION

The EHRUT tested for this study were the various tasks associated with documenting an Implantable Device using the Unique Device Identifier (UDI) and being able to update and retrieve that data in WebHIS 2.0.

Designed to present medical information to healthcare providers in the acute care setting at Holy Name Medical Center, the EHRUT is a clinical information system used by all staff who participate in the care of our patients, on a clinical, financial or record keeping level. The usability testing attempted to represent realistic exercises and conditions.

The purpose of this study was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT). To this end, measures of effectiveness, efficiency, and user satisfaction were captured during the usability testing.

METHOD

PARTICIPANTS

A total of fifteen (15) participants were tested on the EHRUT. Participants in the test were staff nurses, nurses in clinical management, physicians and medical students. Participants work at Holy Name Medical Center in the clinical areas throughout the hospital. Participants did not participate in the requirements gathering, design or development of the EHRUT being tested. Participants were given the

opportunity to have the same orientation and level of training as the actual end users would have received.

Recruited participants had a mix of backgrounds and demographic characteristics conforming to the recruitment screener. The following is a table of participants by characteristics, including demographics, professional experience, computing experience and user needs for assistive technology. Participant names were replaced with Participant IDs so that an individual's data cannot be tied back to individual identities.

	Participant ID	Gender	Age	Education	Occupation/ Role	Professional Experience (mo)	Computer Experience (mo)	Product Experience (mo)	Assistive Technology Needs
1	ID01	Female	30-39	Bachelor's Degree	Staff RN	12	180	12	-
2	ID02	Male	20-29	Bachelor's Degree	Staff RN	60	150	6	-
3	ID03	Female	60-69	Master's Degree	Nursing Director	420	360	120	-
4	ID04	Female	20-29	Bachelor's Degree	Medical Student	0	130	1	-
5	ID05	Female	40-49	Bachelor's Degree	Staff RN	84	240	84	-
6	ID06	Male	30-39	Doctorate Degree	Hospitalist	156	240	36	-
7	ID07	Female	40-49	Associate Degree	Staff RN	240	240	84	-
8	ID08	Male	20-29	Bachelor's Degree	Medical Student	0	180	1	-
9	ID09	Female	40-49	Bachelor's Degree	Staff RN	216	240	72	-
10	ID10	Female	50-59	Master's Degree	Staff RN	420	200	72	-
11	ID11	Male	40-49	Doctorate Degree	Hospitalist	182	192	36	-
12	ID12	Female	60-69	Associate Degree	Operating Room RN	456	120	60	-
13	ID13	Female	30-39	Bachelor's Degree	Staff RN	48	180	24	-
14	ID14	Female	20-29	Bachelor's Degree	Staff RN	36	180	12	-
15	ID15	Female	50-59	Associate Degree	Staff RN	324	240	96	-

Fifteen (15) participants (matching the demographics in the section on Participants) were recruited and fifteen (15) participated in the usability test. Zero (0) participants failed to show for the study.

Participants were scheduled for multiple fifteen (15) and twenty (20) minute sessions with five (5) minutes in between each session for debrief by the administrator(s) and data logger(s), and to reset

systems to proper test conditions where required A spreadsheet was used to keep track of the participant schedule, and included each participant's demographic characteristics.

STUDY DESIGN

The objective of this test was to uncover areas where the application performed well – that is, effectively, efficiently, and with satisfaction – and areas where the application failed to meet the needs of the participants. The data from this test may serve as a baseline for future tests with an updated version of the same EHR and/or comparison with other EHRs provided the same tasks are used. In short, this testing serves as both a means to record or benchmark current usability, but also to identify areas where improvements must be made.

During the usability test, participants interacted with one EHR. Each participant used the system in the same location, and was provided with the same instructions. The system was evaluated for effectiveness, efficiency and satisfaction as defined by measures collected and analyzed for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

Additional information about the various measures can be found in section on Usability Metrics.

TASKS

A number of tasks were constructed that would be realistic and representative of the kinds of activities a user might do with this EHR, including but not limited to:

- ***Documenting an implanted device in a patient's record***
- ***Updating the status of a currently implanted device (removal)***
- ***Reviewing the complete list of devices including all detail***

These are realistic scenarios that are used during procedures and surgeries on a daily basis, whether implanting or removing a said device, implant or prosthetic.

PROCEDURES

Upon arrival, participants were greeted by the test administrator. Their identity was verified and matched with their name on the participant schedule. Participants were then assigned an anonymous participant ID.

To ensure that the test ran smoothly, multiple staff members participated in this test – the administrator and several members of the hospital's simulation center, who are experienced test administrators/loggers and frequently participate in these types of exercises in our busy simulation training and testing center.

The administrator moderated the session including administering instructions and tasks. The administrator in conjunction with the data loggers monitored task times, obtained post-task rating data, and took notes on participant comments, in addition to taking notes on task success, path deviations, number and type of errors and comments.

Participants were instructed to perform the tasks (see specific instructions below):

- As quickly as possible making as few errors and deviations as possible.
- Without assistance; administrators were allowed to give immaterial guidance and clarification on tasks, but not instructions on use.
- Without using a think aloud technique.

For each task, the participants were given a written copy of the task. Task timing began once the administrator finished reading the question. The task time was stopped once the participant indicated they had successfully completed the task. Scoring is discussed below.

Participants' demographic information, task success rate, time on task, errors, deviations, verbal responses, and post-test questionnaire were recorded into a spreadsheet.

TEST LOCATION

The test facility included a quiet computer training room with tables and a desktop computer for each participant. To ensure that the environment was comfortable for users, noise levels were kept to a minimum with the ambient temperature within a normal range. All of the safety instruction and evacuation procedures were valid, in place, reviewed with and visible to the participants.

TEST ENVIRONMENT

The EHRUT would typically be used in a hospital (inpatient unit, procedural area, clinic). In this instance, the testing was conducted in the same facility on computers that were set up the same way as the clinical areas of the hospital where the application will be used. The computer being used was an **IBMi Power 7 running I5OS v7.2.** The environment in which the testing was conducted was the WebHIS 2.0 User Acceptance Testing (UAT) environment.

The participants used a mouse, keyboard and barcode scanner when interacting with the EHRUT. The following is the configuration of the PCs used both in the test environment as well as throughout the organization in the clinical areas where desktop PCs are used:

Windows 10, dual 2.70GHz IntelCore i5 processor
8GB RAM
1600 x 900 resolution
RGB color format, SDR Color Space
6-bit
60 Hz refresh rate
23 inch monitor
Ds8100-HC Series Handheld Imager (for barcode scanning)

The application was set up by the Holy Name Medical Center Information Systems department '**Desktop Support Division'**, according to the vendor's documentation, describing the system set-up and preparation.

System performance (i.e., response time) on the UAT is typically slower than that of the live environment. This was considered when calculating optimal times into the parameters for testing. Additionally, participants were instructed to not change any of the default system settings (such as control of font size).

TEST FORMS AND TOOLS

During the usability test, various documents and instruments were used, including:

- Participant demographic form
- Description of the test cases
- Objective test data capture form (for each test)
- Likert Scale for each test scenario
- System Usability Scale – one overall for each participant

Sample Objective Test Data Forms:

Task Performance Evaluation	**WebHIS 2.0** **Task 175.310(a)14 Implantable Devices**
Task	**Record a new implanted device (UDI) and status**
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	**WebHIS 2.0** **Task 170.315(a)14** **Implantable Devices**
Task	***Change the status of the implanted device from implanted to removed***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

Task Performance Evaluation	***WebHIS 2.0*** ***Task 170.315(a)14 – Implantable Devices***
Task	**Review the list of implanted and removed devices and record**
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

PARTICIPANT INSTRUCTIONS

The administrator reads the following instructions aloud to the each participant:

> *Thank you for participating in this study. Your input is very important. Our Implantable Device List test session today will last for twenty minutes with a five minute break after everyone has completed the exercises. During that time you will use an instance of an electronic health record. I will ask you to complete a few tasks using this system and answer some questions. You should complete the tasks as quickly as possible making as few errors as possible. Please try to complete the tasks on your own following the instructions very closely.* ***Note that you will be using a handheld barcode scanner during this exercise.***
>
> *Please note that we are not testing you, we are testing the system, therefore if you have difficulty with any aspect of the tasks it likely means that something needs to be improved in the system. Please make note of it if you feel it will help us improve the usability of the system.*
>
> *We will be here in case you need general questions, but we are not able to instruct you or provide help in how to use the application. Overall, we are interested in how easy (or how difficult) this system is to use, what in it would be useful to you, and how we could improve it. Please be honest in your opinions. All of the information that you provide will be kept confidential. Your name will not be associated with any test data or comments at any time. Should you feel it necessary you are able to withdraw at any time during the testing.*

Following the procedural instructions, participants were shown the EHR and as their first task, were given 10 minutes to explore the system and ask general questions. Once this task was complete, the administrator gave the following instructions:

> *For each task, I will read the description to you and say "Begin." At that point, please perform the task and say "Done" once you believe you have successfully completed the task. I would like to request that you not talk aloud or verbalize while you are doing the tasks. I will ask you your impressions about the task once you are done.*

Participants were then given three (3) tasks within the Patient Implantable Device Maintenance program to complete and evaluate.

USABILITY METRICS

According to the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records*, EHRs should support a process that provides a high level of usability for all users. The goal is for users to interact with the system effectively, efficiently, and with an acceptable level of satisfaction. To this end, metrics for effectiveness, efficiency and user satisfaction were captured during the usability testing. The goals of the test were to assess:

1. Effectiveness of WebHIS 2.0 by measuring participant success rates and errors
2. Efficiency of WebHIS 2.0 by measuring the average task time and path deviations
3. Satisfaction with WebHIS 2.0 by measuring usability, using the System Usability Scale (SUS) overall (at the end of all the testing) and a Likert scale for ease of use for each exercise.

DATA SCORING

The following table details how tasks were scored, errors evaluated, and the time data analyzed.

Measures	Rationale and Scoring
Effectiveness: Task Success	A task was counted as a "Success" if the participant was able to achieve the correct outcome, without assistance, within the time allotted on a per task basis. The total number of successes were calculated for each task and then divided by the total number of times that task was attempted. The results are provided as a percentage. Task times were recorded for successes. Observed task times divided by the optimal time for each task is a measure of optimal efficiency. Optimal task performance time, as benchmarked by expert performance under realistic conditions, is recorded when constructing tasks. Target task times used for task times in the Moderator's Guide must be operationally defined by taking multiple measures of optimal performance and multiplying by some factor [e.g., 1.25] that allows some time buffer because the participants are presumably not trained to expert performance. Thus, if expert, optimal performance on a task was [x] seconds then allotted task time performance was [x * 1.25] seconds. This ratio should be aggregated across tasks and reported with mean and variance scores.
Effectiveness: Task Failures	If the participant abandoned the task, did not reach the correct answer or performed it incorrectly, or reached the end of the allotted time before successful completion, the task was counted as a "Failure." No task times were taken for errors. The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant. On a qualitative level, an enumeration of errors and error types should be collected.
Efficiency: Task Deviations	The participant's path (i.e., steps) through the application was recorded. Deviations occur if the participant, for example, went to a wrong screen, clicked on an incorrect menu item, followed an incorrect link, or interacted incorrectly with an on-screen control. This path was compared to the optimal path. The number of steps in the observed path is divided by the number of optimal steps to provide a ratio of path deviation. It is strongly recommended that task deviations be reported. Optimal paths (i.e., procedural steps) should be recorded when constructing tasks.
Efficiency: Task Time	Each task was timed from when the administrator said "Begin" until the participant said, "Done." If he or she failed to say "Done," the time was stopped when the participant stopped performing the task. Only task times for tasks that were successfully completed were included in the average task time analysis. Average time per task was calculated for each task. Variance measures (standard deviation and standard error) were also calculated.
Satisfaction: Task Rating	Participant's subjective impression of the ease of use of the application was measured by administering both a simple post-task question as well as a post-session questionnaire. After each task, the participant was asked to rate "Overall, this task was:" on a scale of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants. Common convention is that average ratings for systems judged easy to use should be 3.3 or above. To measure participants' confidence in and likeability of the [EHRUT] overall, the testing team administered the System Usability Scale (SUS) post-test questionnaire. Questions included, "I think I would like to use this system frequently," "I thought the system was easy to use," and "I would imagine that most people would learn to use this system very quickly." See full System Usability Score questionnaire in Appendix.

RESULTS

DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above. Participants who failed to follow session and task instructions had their data excluded from the analyses.

The usability testing results for the EHRUT are detailed below. The results should be seen in light of the objectives and goals outlined in the Study Design section. The data should yield actionable results that if corrected, yield material, positive impact on user performance.

Measure 170.315(g)3	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
(a)14 Task	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Record implanted device	15	100%	1.26	47 sec	78.33%	1.53	3.73
Update removal of implanted device	15	100%	1.11	17.73 sec	70.93%	0.33	4.33
Review active and removed devices	15	100%	1.17	19.07 sec	76.27%	0.67	4.27

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 86%. Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80% would be considered above average.

MAJOR FINDINGS

- Major Findings –
  - Initial screen labels are not obvious to a new user – example – buttons for 'UDI's' and 'Free Text'
  - The 'Add' button was initially difficult to find
  - Prompts for scanning did not indicate which barcodes to scan
  - After scanning the barcodes, there was no indication whether or not complete data was obtained; users stated they were able to save incomplete data
  - The 'remove' button should be a contrasting color
  - The full loop from insertion through removal (if applicable) worked well and provided a clear audit trail for each implant, including site, notes, status updates and reason for removal, which is a mandatory field.
  - The detail screen is too large for the display area and the user must scroll down to see the Save button.

AREAS FOR IMPROVEMENT

- Areas for Improvement
  - Include 'tool tips' over each button
  - Move the 'Add' button to the left side of the screen where the other options are located
  - Colorize the 'Add' button in a color other than the rest of the buttons (gray)
  - Allow the user to scan additional barcodes and indicate when all necessary data is obtained and documented (create additional fields)
  - Highlight the 'remove' button to make it more obvious to the user
  - The system will indicate when it has captured all data, providing the user with the ability to decide whether or not to save the item
  - Condensing the 'add' screen to avoid scrolling to the Save button.

EFFECTIVENESS

Overall it was found that the effectiveness of the Implantable device tracking program was high. Although users did make valuable observations and recommendations regarding the User Interface, the application itself captures and displays current, inactive and removed implants in a clear manner. Based on this easy to use application, potential recalls would be very easy to manage, and requested reports will be easy to generate.

EFFICIENCY

The efficiency of this application based on time

Enter new implant – the optimal time for this task was determined to be 60 seconds. The average time as demonstrated during the test was 47 seconds.

Modify the implant status – the optimal time for this task was determined to be 25 seconds. The average time as demonstrated during the test was 18 seconds. (17.7)

Review implant data – the optimal time for this task was determined to be 25 seconds. The average time as demonstrated during the test was 19 seconds.

Efficiency of this application based on path deviations (number of clicks):

Enter new implant – the optimal number of clicks for this task was determined to be 6. The average number of clicks demonstrated during the test was 7.5.

Change the implant status – the optimal number of clicks for this task was determined to be 3. The average number of clicks demonstrated during this test was 3.3

Review implant data – The optimal number of clicks for this task was determined to be 4. The average number of clicks demonstrated during this

SATISFACTION

Utilizing a Likert Scale of 1-5, 5 being 'very easy' to 1 being 'very difficult', the average user rating for each of the Clinical Decision Support Rules was as follows:

Adding new implant (UDI)	**3.7**
Change the status of implant	**3.9**
Review patient history for all implants	**4.0**

PARTICIPANT COMMENTS

Participant comments overall reflected a high satisfaction rate with the Implantable Device adding and editing screens, as noted in the following quotes collected from the follow up forms:

> "Easy to collect and track important data".
> "I was easily able to add an implant to my patient's record".
> "It is reassuring to know that we can retrieve this information whenever necessary".
> "This is a game-changer for those of us who work in the Operating Room where we implant devices on a daily basis. Thank you!"

EHR Usability Test Report – 170.315(g)(3) Safety Enhanced Design CCDA Reconciliation (b)2

UCD Process Used: NISTIR 7741:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7741) NIST, Guide to the Processes Approach for Improving the Usability of Electronic Health Records. NIST, 29 Nov. 2010. Web. https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7741.pdf

Report Based on: NISTIR 7742:

Schumacher, Robert M., and Svetlana Z. Lowry. (NISTIR 7742) NIST, Customized Common Industry Format Template for Electronic Health Record Usability Testing. NIST, 16 Nov. 2010. Web.
https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7742.pdf

Contents

EHR USABILITY TEST REPORT – CCDA Import and Reconciliation (b)2

Report based on ISO/IEC 25062:2006 Common Industry Format for Usability Test Reports

Date of Usability Test: July 16, 2019
Date Usability Test was Conducted: July 16, 2019
Date Report was Prepared: July 22, 2019
Report Prepared by: Deborah Ross
Name of System Test Laboratory (STL): Holy Name Medical Center
STL Contact Person, Title and Affiliation: Deborah Ross, Clinical Analyst
STL Phone Number: 201-833-7114
STL Email Address: debross@mail.holyname.org
STL Mailing Address: 718 Teaneck Road, Teaneck, New Jersey 07666

EXECUTIVE SUMMARY

A usability test of the CCDA reconciliation program in WebHIS 2.0 was conducted on July 16, 2019, in Teaneck, New Jersey by Holy Name Medical Center. The purpose of this test was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT).

During the usability test, ten (10) healthcare providers matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative tasks. All of the participants were members of the Health Information Management Department, serving in various roles.

The study collected performance data and user reactions to multiple tasks related to matching and reconciling a patient's Continuity of Care Document (CCD) from an outside provider into the native WebHIS application.

- Access a Continuity of Care Document
- Identify and match to a patient
- Review and incorporate relevant data into the WebHIS 2.0 application

During the twenty (20) minute summative usability test, each participant was greeted by the administrator and instructed that they could withdraw from the exercise at any time. All participants had varying levels of prior experience with the EHR. The administrator introduced the test, and instructed participants to complete a series of tasks (given one at a time) using the EHRUT. During the testing, the administrator timed the test, and along with the logger(s) recorded user performance data on paper. The administrator did not give the participant assistance in how to complete the task.

The following types of data were collected for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations
- Participant's satisfaction ratings of the system – LIKERT and System Usability Scale (SUS)

All participant data was de-identified – no correspondence could be made from the identity of the participant to the data collected. Various recommended metrics, in accordance with the examples set forth in the NIST Guide to the Processes Approach for Improving the Usability of Electronic Health

Records were used to evaluate the usability of the EHRUT. The summary data collected for *170.315(b)2* is listed in the **RESULTS** section of this document. The following is the recommended template utilized when calculating the individual results:

Measure 170.315(g)3	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
(b)2 Task	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Retrieve, match and incorporate CCD data	10						

The overall results calculated from the **System Usability Scale (SUS)** provided to each participant scored the subjective satisfaction with the system based on performance with these tasks to be **89%.** This scale was provided to each participant **only one time at the end of all of the tests**.

In addition to the performance data, qualitative data including both major findings and areas for improvement, based on participant anecdotal feedback will be listed in the RESULTS section as well.

INTRODUCTION

The EHRUT tested for this study were the various tasks associated with recording, updating and reviewing a patient's Medication Allergies in WebHIS 2.0.

Designed to present medical information to healthcare providers in the acute care setting at Holy Name Medical Center, the EHRUT is a clinical information system used by all staff who participate in the care of our patients, on a clinical, financial or record keeping level. The usability testing attempted to represent realistic exercises and conditions.

The purpose of this study was to test and validate the usability of the current user interface, and provide evidence of usability in the EHR Under Test (EHRUT). To this end, measures of effectiveness, efficiency, and user satisfaction were captured during the usability testing.

METHOD

PARTICIPANTS

A total of ten (10) participants were tested on this application in the EHRUT. Participants in the test were various members of the Health Information Management (HIM) team, including Clinical Documentation Specialists, Coders and Chart Abstracters. Participants work at Holy Name Medical Center. Participants did not participate in the requirements gathering, design or development of the EHRUT being tested. Participants were given the opportunity to have the same orientation and level of training as the actual end users would have received.

Recruited participants had a mix of backgrounds and demographic characteristics conforming to the recruitment screener. The following is a table of participants by characteristics, including demographics, professional experience, computing experience and user needs for assistive technology. Participant names were replaced with Participant IDs so that an individual's data cannot be tied back to individual identities.

	Participant ID	Gender	Age	Education	Occupation/ Role	Professional Experience (mo)	Computer Experience (mo)	Product Experience (mo)	Assistive Technology Needs
26	ID01	F	40-49	High School graduate,diploma or the equivalent	Medical Records Q/A	240	240	156	N
27	ID02	F	60-69	Bachelor's Degree	Supervisor	300	300	120	N
28	ID03	F	60-69	Trade/Technical/Vocational Training	HIM Associate	336	336	156	N
29	ID04	F	40-49	Associate Degree	HIM Associate	324	324	156	N
30	ID05	F	60-69	Master's Degree	Clinical Documentation Supervisor	444	240	108	N
31	ID06	F	50-59	Bachelor's Degree	Clinical Documentation Specialist	372	324	156	N
32	ID07	F	30-39	Associate Degree	Clinical Documentation Specialist	228	240	156	N
33	ID08	F	30-39	Associate Degree	Clinical Documentation Specialist	204	300	156	N
34	ID09	F	60-69	Trade/Technical/Vocational Training	HIM Associate	480	240	84	N
35	ID10	M	40-49	Associate Degree	Outpatient Coder	48	360	48	N

Ten (10) participants (matching the demographics in the section on Participants) were recruited and ten (10) participated in the usability test. Zero (0) participants failed to show for the study.

Participants were scheduled for a thirty (30) minute session. A spreadsheet was used to keep track of the participant schedule, and included each participant's demographic characteristics.

STUDY DESIGN

The objective of this test was to uncover areas where the application performed well – that is, effectively, efficiently, and with satisfaction – and areas where the application failed to meet the needs of the participants. The data from this test may serve as a baseline for future tests with an updated version of the same EHR and/or comparison with other EHRs provided the same tasks are used. In short,

this testing serves as both a means to record or benchmark current usability, but also to identify areas where improvements must be made.

During the usability test, participants interacted with one EHR. Each participant used the system in the same location, and was provided with the same instructions. The system was evaluated for effectiveness, efficiency and satisfaction as defined by measures collected and analyzed for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

Additional information about the various measures can be found in the section on Usability Metrics.

TASKS

A composite task was designed that would be realistic and representative of the kinds of activities a user might do with this EHR, which included the following components:

- ***Retrieve a CCDA from an outside practice***
- ***Match the patient demographics to that of a patient already in the native EMR (WebHIS)***
- ***Review and incorporate clinical data from the CCDA into the native EMR (WebHIS)***

These are components of a realistic scenario that would occur on a daily basis beginning with the staff of the Health Information Management department.

PROCEDURES

Upon arrival, participants were greeted by the test administrator. Their identity was verified and matched with their name on the participant schedule. Participants were then assigned an anonymous participant ID.

This test was conducted on an individual basis – each participant performed the functions individually, with the administrator both describing and conducting the test.

The administrator moderated the session including administering instructions and tasks. The administrator monitored task times, obtained post-task rating data, and took notes on participant comments, in addition to taking notes on task success, path deviations, number and type of errors and comments.

Participants were instructed to perform the tasks (see specific instructions below):

- As quickly as possible making as few errors and deviations as possible.
- Without assistance; administrators were allowed to give immaterial guidance and clarification on tasks, but not instructions on use.
- Without using a think aloud technique.

For each task, the participants were given a written copy of the task. Task timing began once the administrator finished reading the question. The task time was stopped once the participant indicated they had successfully completed the task. Scoring is discussed below.

Participants' demographic information, task success rate, time on task, errors, deviations, verbal responses, and post-test questionnaire were recorded into a spreadsheet.

TEST LOCATION

The test facility included a quiet computer training room with tables and a desktop computer for each participant. To ensure that the environment was comfortable for users, noise levels were kept to a minimum with the ambient temperature within a normal range. All of the safety instruction and evacuation procedures were valid, in place, reviewed with and visible to the participants.

TEST ENVIRONMENT

The EHRUT would typically be used in a hospital HIM department. In this instance, the testing was conducted in the same facility on computers that were set up the same way as the clinical areas of the hospital where the application will be used. The computer being used was an **IBMi Power 7 running I5OS v7.2.** The environment in which the testing was conducted was the WebHIS 2.0 User Acceptance Testing (UAT) environment.

The participants used a mouse and keyboard when interacting with the EHRUT. The following is the configuration of the PCs used both in the test environment as well as throughout the organization in the clinical areas where desktop PCs are used:

Windows 10, dual 2.70GHz IntelCore i5 processor
8GB RAM
1600 x 900 resolution
RGB color format, SDR Color Space
6-bit
60 Hz refresh rate
23 inch monitor

The application was set up by the Holy Name Medical Center Information Systems department '**Desktop Support Division'**, according to the vendor's documentation, describing the system set-up and preparation. System performance (i.e., response time) on the UAT is typically slower than that of the live environment. This was considered when calculating optimal times into the parameters for testing. Additionally, participants were instructed to not change any of the default system settings (such as control of font size).

TEST FORMS AND TOOLS

During the usability test, various documents and instruments were used, including:

- Participant demographic form
- Description of the test cases
- Objective test data capture form (for each test)
- Likert Scale for each test scenario

- System Usability Scale – one overall for each participant

Sample Objective Test Data Forms:

Task Performance Evaluation	**WebHIS 2.0** **Task 175.310(b)2 CCDA**
Task	***Match, reconcile and incorporate a patient's CCD for Allergies, Meds and Problems***
Participant number	
Task Time (minutes:seconds)	
Task Success	☐ Easily Completed ☐ Completed with difficulty ☐ Not able to complete
Path Deviation	☐ Yes ☐ No
Observed number of steps	
Number of errors	
Participants Verbalizations	
Major Findings	
Satisfaction Rating	☐ 1 Very Difficult ☐ 2 ☐ 3 ☐ 4 ☐ 5 Very Easy

PARTICIPANT INSTRUCTIONS

The administrator reads the following instructions aloud to the each participant:

> *Thank you for participating in this study. Your input is very important. The Clinical Information Reconciliation and Incorporation session today will last for twenty minutes. During that time you will use an instance of an electronic health record. I will ask you to complete a few tasks using this system and answer some questions. You should complete the tasks as quickly as possible making as few errors as possible. Please try to complete the tasks on your own following the instructions very closely.*
>
> *Please note that we are not testing you, we are testing the system, therefore if you have difficulty with any aspect of the tasks it likely means that something needs to be improved in the system. Please make note of it if you feel it will help us improve the usability of the system.*
>
> *We will be here in case you need general questions, but we are not able to instruct you or provide help in how to use the application. Overall, we are interested in how easy (or how difficult) this system is to use, what in it would be useful to you, and how we could improve it. Please be honest in your opinions. All of the information that you provide will be kept confidential. Your name will not be associated with any test data or comments at any time. Should you feel it necessary you are able to withdraw at any time during the testing.*

Following the procedural instructions, participants were shown the EHR and as their first task, were given 10 minutes to explore the system and ask general questions. Once this task was complete, the administrator gave the following instructions:

> *For each task, I will read the description to you and say "Begin." At that point, please perform the task and say "Done" once you believe you have successfully completed the task. I would like to request that you not talk aloud or verbalize while you are doing the tasks. I will ask you your impressions about the task once you are done.*

Participants were then given instructions on Reconciling and Incorporating a patient's clinical data into their record in the EHRUT.

USABILITY METRICS

According to the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records*, EHRs should support a process that provides a high level of usability for all users. The goal is for users to interact with the system effectively, efficiently, and with an acceptable level of satisfaction. To this end, metrics for effectiveness, efficiency and user satisfaction were captured during the usability testing. The goals of the test were to assess:

1. Effectiveness of WebHIS 2.0 by measuring participant success rates and errors
2. Efficiency of WebHIS 2.0 by measuring the average task time and path deviations
3. Satisfaction with WebHIS 2.0 by measuring usability, using the System Usability Scale (SUS) overall (at the end of all the testing) and a Likert scale for ease of use for each exercise.

DATA SCORING

The following table details how tasks were scored, errors evaluated, and the time data analyzed.

Measures	Rationale and Scoring
Effectiveness: Task Success	A task was counted as a "Success" if the participant was able to achieve the correct outcome, without assistance, within the time allotted on a per task basis. The total number of successes were calculated for each task and then divided by the total number of times that task was attempted. The results are provided as a percentage. Task times were recorded for successes. Observed task times divided by the optimal time for each task is a measure of optimal efficiency. Optimal task performance time, as benchmarked by expert performance under realistic conditions, is recorded when constructing tasks. Target task times used for task times in the Moderator's Guide must be operationally defined by taking multiple measures of optimal performance and multiplying by some factor [e.g., 1.25] that allows some time buffer because the participants are presumably not trained to expert performance. Thus, if expert, optimal performance on a task was [x] seconds then allotted task time performance was [x * 1.25] seconds. This ratio should be aggregated across tasks and reported with mean and variance scores.
Effectiveness: Task Failures	If the participant abandoned the task, did not reach the correct answer or performed it incorrectly, or reached the end of the allotted time before successful completion, the task was counted as a "Failure." No task times were taken for errors. The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant. On a qualitative level, an enumeration of errors and error types should be collected.
Efficiency: Task Deviations	The participant's path (i.e., steps) through the application was recorded. Deviations occur if the participant, for example, went to a wrong screen, clicked on an incorrect menu item, followed an incorrect link, or interacted incorrectly with an on-screen control. This path was compared to the optimal path. The number of steps in the observed path is divided by the number of optimal steps to provide a ratio of path deviation. It is strongly recommended that task deviations be reported. Optimal paths (i.e., procedural steps) should be recorded when constructing tasks.
Efficiency: Task Time	Each task was timed from when the administrator said "Begin" until the participant said, "Done." If he or she failed to say "Done," the time was stopped when the participant stopped performing the task. Only task times for tasks that were successfully completed were included in the average task time analysis. Average time per task was calculated for each task. Variance measures (standard deviation and standard error) were also calculated.
Satisfaction: Task Rating	Participant's subjective impression of the ease of use of the application was measured by administering both a simple post-task question as well as a post-session questionnaire. After each task, the participant was asked to rate "Overall, this task was:" on a scale of 1 (Very Difficult) to 5 (Very Easy). These data are averaged across participants. Common convention is that average ratings for systems judged easy to use should be 3.3 or above. To measure participants' confidence in and likeability of the [EHRUT] overall, the testing team administered the System Usability Scale (SUS) post-test questionnaire. Questions included, "I think I would like to use this system frequently," "I thought the system was easy to use," and "I would imagine that most people would learn to use this system very quickly." See full System Usability Score questionnaire in Appendix.

RESULTS

DATA ANALYSIS AND REPORTING

The results of the usability test were calculated according to the methods specified in the Usability Metrics section above. Participants who failed to follow session and task instructions had their data excluded from the analyses.

The usability testing results for the EHRUT are detailed below. The results should be seen in light of the objectives and goals outlined in the Study Design section. The data should yield actionable results that if corrected, yield material, positive impact on user performance.

Measure 170.315(g)3	N	Task Success	Path Deviation	Task Time		Errors	Task Ratings (5=Easy)
(b)2	#	Mean (SD)	Deviations (Observed/Optimal)	Mean (SD) (sec)	Deviations (Observed/Optimal)	Mean (SD)	Mean (SD)
Reconcile and incorporate clinical information from a CCD	**10**	**100%**	**1.02**	**51.1 sec**	**68.1%**	**0.20**	**4.40**

The results from the SUS (System Usability Scale) scored the subjective satisfaction with the system based on performance with these tasks to be 86%. Broadly interpreted, scores under 60 represent systems with poor usability; scores over 80 would be considered above average.

MAJOR FINDINGS

- Overall the system was easy to use and will provide significant value when receiving patients from other facilities.
- There is no link to the patient's record in the WebHIS 2.0 for additional demographic and clinical verifications.
- The system is intuitive to use.
- The CCD screen displaying the various categories of clinical data should distinguish among the categories for easier viewing.
- It is easy to distinguish the clinical information that was imported from the CCD from clinical information native to the WebHIS 2.0. in the patient record.
- The link to the CCD is not labeled.

AREAS FOR IMPROVEMENT

- List the patient name in the search as a hyperlink so that the user can validate additional demographic information without leaving the screen.
- Distinguish among the various clinical categories (allergies, medications, problems, etc) on the screen to help the user target the specific data that they are looking for.
- Include a label next to the button that confirms the patient match. Currently it is simply a checkbox at the end of the row of demographic information.

- Make this same checkbox larger than it currently is.

EFFECTIVENESS

The Clinical Information Reconciliation and Incorporation program is a useful, effective application. It provides a streamlined way of accessing external records, matching them to current patients and allowing the staff to review the discrete data elements and determine whether or not they should be imported into the WebHIS 2.0.

EFFICIENCY

The efficiency of this application based on time –

The optimal time for completing this exercise was determined to be 75 seconds. The average time as demonstrated by the 10 participants was calculated to be 51 seconds. This demonstrates the usability and intuitive nature of these screens. Standard design conventions are employed in the design of all screens and applications.

Efficiency of this application based on path deviations (number of clicks) -

Based on the optimal number of 'clicks' calculated to complete this exercise, only 2 users deviated from this path, and then only by a single click, demonstrating the efficient, intuitive design of the application.

SATISFACTION

Utilizing a Likert Scale of 1-5, 5 being 'very easy' to 1 being 'very difficult', the average user rating for the Clinical Information Reconciliation and Incorporation program was -

Clinical Information Reconciliation and Incorporation - 4.40

PARTICIPANT COMMENTS

Participant comments overall reflected a high satisfaction rate with the Clinical Information Reconciliation and Incorporation application. The staff of HIM was very pleased to see such an easy intuitive application designed for them, as noted in the following quotes collected from the participants:

"I can see this being very useful, especially since I am involved with reviewing records that come in for patients who are transferred from other organizations".

"Easy and straight forward"

"The tasks were very easy and worked very well. It was very clear and easy to understand"

"This will be a positive enhancement to our current, manual workflow"

WebHIS v1.2

2024 EHR Usability Test

(b)(11) Addendum Report

Safety Enhanced Design 170.315 (g}{3} - October 2024

Report based on ISO/IEC 62366 Common Industry Format for Usability Test Reports

WebHIS - Safety Enhanced Design

Date of Usability Test: Started September 4, 2024 - Concluded October 27, 2024

Date of Report: November 7, 2024

Report Prepared By: EMR Advocate

Main Contact: Anitha Edward
Phone: 201-833-3000 Ext: 3507
www.holyname.org
aedward@holyname.org
718 Teaneck Road
Teaneck, NJ 07666

Table of Contents

Executive Summary

Usability tests of the 1.2 version of the EHR were conducted at various times during the development cycle, the last session for which was held on November 7th, 2024. The purpose of these tests was to test and validate the usability of the current user interface, and provide evidence of usability of the EHR Under Test (EHRUT).

During the usability test, a combination of test participants and clinicians matching the target demographic criteria served as participants and used the EHRUT in simulated, but representative tasks.

This study collected performance data on 14 tasks typically conducted in the EHR:

Decision Support Intervention (Evidence Based and User-supplied Predictive)

- Configuration/enablement
- Source attribute management record and change
- DSI Selection and access
- Feedback loop entries and report export (Evidence Based Only)

Tasks were selected based on their frequency of use, criticality of function, and those that may be most troublesome for users. Tasks were constructed in light of the study objectives. A detailed list of the tasks provided to the participants can be accessed from Appendix B.

During the 65-minute, one-on-one, remote usability test, each participant was greeted by the. Participants were then assigned a participant ID and asked to review and sign an informed consent/release form. Participants were read an overview of the test, its intended purpose, general instructions, and then advised that they could withdraw at any time. Participants had no prior experience with the EHR.

The administrator introduced the test, and instructed the participant to complete a series of tasks (given one at a time) using the EHRUT. During the testing, the administrator timed the test and, along with the data logger(s) recorded user performance data on paper and electronically. The administrator did not give the participant assistance in how to complete the task.

The test session, including participant screens, user workflow, and audio, was recorded for subsequent analysis.

The following types of data were collected for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbal feedback
- Participant's task effort ratings of the system using a Likert Scale

All participant data was de-identified so that no correlation could be made from the identity of the participant to the data collected. Following the conclusion of the testing, participants were asked to complete a post-test questionnaire. Participants were not compensated for their time.

Various recommended metrics, in accordance with the examples set forth in the NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records, were used to evaluate the usability of the EHRUT. Following is a summary of the performance and rating data collected on the EHRUT.

Introduction

This study is the result of usability testing performed on the 1.2 version of the EHR, which is designed to collect, track, and report medical information collected from healthcare providers in an ambulatory setting. The application consists of solutions for a range of services including medical, dental, vision, and behavior allowing practices to provide patient care for all their services.

The usability testing attempted to represent realistic exercises and conditions. The purpose of this study was to test and validate the usability of the current user interface, and provide evidence of usability to support certification according to criteria outlined in Safety Enhanced Design §170.31S(g){3), specifically:

- § 170.315 (b)(11) Clinical decision support - Evidence Based
- § 170.315 (b)(11) Clinical decision support – User-supplied Predictive

Method

Participants

A total of ten (10) participants were tested on the EHR. Participants in the test included doctors, medical assistants, clinic managers, and test participants general office aptitude for technology. Volunteer participants were recruited by and were not compensated for their time.

Participants had no direct connection to the development of or organization producing the EHR, and they were not from or affiliated with , and did not need any orientation or training as they all were experienced EHR users.

For test purposes, end-user characteristics were identified and translated into a recruitment screener used to solicit potential participants.

Participants had a mix of backgrounds and demographic characteristics. The following is a table of participants by characteristics, including demographics, professional experience, computing experience, and user needs for assistive technology. Participant names were replaced with Participant IDs so that an individual's data cannot be tied back to their identity.

User ID	Sex	Age	Education	Occupation/Role	Professional Experience (Months)	Computer Experience (months)	Product Experience (Months)	Assistive Technology
1	Male	60-69	Doctorate degree	MD - Family Medicine	240	200	0	No
2	Female	40-49	Masters degree	Health IT Consultant	192	120	0	No
3	Female	20-29	Some college credit, no degree	Front Office Administrator	168	136	0	No
4	Male	30-39	Bachelors degree	Registered Nurse	132	264	0	No
5	Female	40-49	Bachelors degree	Healthcare Policy Analyst	180	120	0	No
6	Male	40-49	Masters Degree	Physician Assistant	204	204	0	No
7	Female	60-69	Doctorate degree	Physician/ Physiatry	240	228	0	No
8	Female	30-39	Associates degree	Medical Assistant	156	120	0	No
9	Male	20-29	Associates degree	Medical Assistant	102	96	0	No
10	Male	50-59	Doctorate degree	Clinical Psychologist	168	150	0	No

10 participants participated in the usability test. 0 participants failed to show for the study.

Participants were scheduled for 65-minute sessions with 5 minutes in between each session for debrief by the administrator and data logger, and to reset systems to proper test conditions. A spreadsheet was used to keep track of the participant schedule, and included each participant's demographic characteristics as provided by the participant.

Study Design

Overall, the objective of this test was to uncover areas where the application performed well – that is, effectively, efficiently, and with satisfaction – and areas where the application failed to meet the needs of the participants. The data from this test may serve as a baseline for future tests with an updated version of the same EHR and/or comparison with other EHRs provided the same tasks are used. In short, this testing serves as both a means to record or benchmark current usability, but also to identify areas where improvements must be made.

During the usability test, participants interacted with one EHR. Each participant used the system in the same development environment, and was provided with the same instructions. The system was evaluated for effectiveness, efficiency and satisfaction as defined by measures collected and analyzed for each participant:

- Number of tasks successfully completed within the allotted time without assistance
- Time to complete the tasks
- Number and types of errors
- Path deviations
- Participant's verbalizations (comments)
- Participant's satisfaction ratings of the system

Additional information about the various measures can be found in the Section on Usability Metrics.

Tasks

In support certification according to criteria outlined in Safety Enhanced Design §170.315(g)(3), 14 tasks were constructed that would be realistic and representative of the kinds of activities a user might conduct with the EHR, in the following overall categories:

- Decision Support Intervention - Evidence Based
- Decision Support Intervention – User-supplied Predictive

Tasks were selected based on their frequency of use, criticality of function, and those that may be most troublesome for users. Tasks were designed to meet the study objectives. A detailed list of the tasks provided is included in Appendix B.

Procedures

Remote testing was conducted via a Zoom session by a proctor with 10+ years' experience with the EHRUT. A Remote testing methodology was selected to both for convenience to accommodate the volunteer participants but also because that technology includes recording of the screen-sharing and audio for subsequent review and analysis.

Participants were advised to choose a quiet location to participate in the study using their own computers, and to:

- Complete the tasks as quickly as possible, using their normal workflow
- Complete the tasks without assistance except to clarify task details, if necessary

All test sessions were recorded by Zoom and subsequently analyzed. While participants completed the tasks, an observer monitored task times, obtained post-task rating data, and took notes on participant comments, and the data logger and took notes on task success, path deviations, number and type of errors, and comments.

Participants' demographic information, task success rate, time on task, errors, deviations, verbal responses, and post test questionnaire were recorded into a spreadsheet. Participants were thanked for their time.

Test Location

Test sessions were conducted remotely via a Zoom meeting. The test administrator, observers, and participant logged into the session from their various locations. All observers and the data logger could see the participant's screen, and listen to the audio of the session.

Test Environment

The EHRUT would be typically be used in a healthcare office or facility. In this instance, the testing was conducted remotely via a Zoom meeting from the participants location origin. For testing, the proctor hosted the EHRUT as a Microsoft Remote Desktop Application running on Windows Server 2016.

The participants used their own hardware including; computer, keyboard, and mouse when testing.

Test Forms and Tools

During the usability test, various documents and instruments were used, including:

- Proctor Guide
- Participant Guide

The Proctor's Guide was devised to be able to capture required data. The participant's interaction with the EHR application was captured and recorded via the Zoom meeting technology.

Participant Instructions

The proctor read the following instructions to each participant:

Thank you for participating in this study. Your input is very important. Our session today will last about 65 minutes. During this time, you will be using the current version of the EHR. I will ask you to complete a few tasks using this system and answer some questions. You should complete the tasks as quickly as possible, making as few errors as possible. Please try to complete the tasks on your own following the instructions very closely. Please note that we are not testing you, rather, we are testing the system.
Therefore, if you have difficulty all this means is that something needs to be improved in the system. I will be here in case you need specific help, but I am not able to instruct you or provide help in how to use the application.

Overall, we are interested in how easy (or possibly how difficult) this system is to use, what in it would be useful to you, and how we could improve it.

> *Please be honest with your opinions. All the information that you provide will be kept confidential and your name will not be associated with your comments at any time.*
> *Should you feel it necessary, you are able to withdraw at any time during the testing.*

Following the procedural instructions, participants were logged into the EHRUT and then given tasks to complete based on their role, and the administrator gave the following instructions:

> *For each task, I will read the description to you and say, "Begin.,, At that point, please perform the task and say, "Done,,, once you believe you have successfully completed the task. I will ask you your impressions about the task once you are done.*

Participants were then given their tasks to complete.

Usability Metrics

According to the *NIST Guide to the Processes Approach for Improving the Usability of Electronic Health Records,* EHRs should support a process that provides a high level of usability for all users. The goal is for users to interact with the system effectively, efficiently, and with an acceptable level of satisfaction. To this end, metrics for effectiveness, efficiency and user satisfaction were captured during the usability testing. The goals of the test were to assess:

- Effectiveness of WebHIS v1.2 by measuring participant success rates and errors
- Efficiency of WebHIS v1.2 by measuring the average task time and path deviations
- Satisfaction with WebHIS v1.2 by measuring ease of use ratings

Data Scoring

The following table details how tasks were scored, errors evaluated, and the time data analyzed.

Measures	Rationale and Scoring
Effectiveness: Task Success	A task was counted as a "Success" if the participant was able to achieve the correct outcome, without assistance, within the time allotted on a per task basis. The total number of successes were calculated for each task and then divided by the total number of times that task was attempted. The results are provided as a percentage. Task times were recorded for successes. Observed task times divided by the optimal time for each task is a measure of optimal efficiency. Optimal task performance time, as benchmarked by expert performance under realistic conditions, is recorded when constructing tasks.
Effectiveness: Task Failures	If the participant abandoned the task, did not reach the correct answer or performed it incorrectly, or reached the end of the allotted time before successful completion, the task was counted as an "Failures." No task times were taken for errors. The total number of errors was calculated for each task and then divided by the total number of times that task was attempted. Not all deviations would be counted as errors. This should also be expressed as the mean number of failed tasks per participant. On a qualitative level, an enumeration of errors and error types should be collected.

Measures	Rationale and Scoring
Efficiency: Task Deviations	The participant's path, i.e., steps through the application, was recorded. Deviations occur if the participant, for example, went to a wrong screen, clicked on an incorrect menu item, followed an incorrect link, or interacted incorrectly with an on-screen control. This path was compared to the optimal path. The number of steps in the observed path is divided by the number of optimal steps to provide a ratio of path deviation. It is strongly recommended that task deviations be reported. Optimal paths (i.e., procedural steps) should be recorded when constructing tasks
Efficiency: Task Time	Each task was timed from when the administrator said "Begin" until the participant said, "Done." If he or she failed to say "Done," the time was stopped when the participant stopped performing the task. Only task times for tasks that were successfully completed were included in the average task time analysis. Average time per task was calculated for each task. Variance measures (standard deviation and standard error) were also calculated.

Measures	Rationale and Scoring
Satisfaction: Task Rating	Each participant's subjective impression of the ease-of-use of the application was measured by administering a simple post-task question. After each task, the participant was asked to rate "Overall, this task was easy:" on a scale of 1 (Strongly Agree) to 5 (Strongly Disagree). This data was averaged across participants. Common convention is that average ratings for systems judged easy-to-use should be 3.3 or below. To measure participants' confidence in and likeability of Patient Pattern overall, the testing team administered using a verbal confirmation of the Likert ranking.

Risk (Pre-test)

Before conducting the usability testing for the designated capabilities within the Certified Electronic Health Record Technology (CEHRT), it is essential to assess the pre-test risks associated with each task. This risk assessment will help identify potential user safety concerns and usability issues that may arise during the testing process.

The pre-test risk assessment will consider factors such as the complexity of the tasks, potential for user error, and the impact of any identified risks on patient safety and care quality. By evaluating these risks, we can implement appropriate mitigation strategies to enhance the effectiveness of the user-centered design (UCD) processes.

Below is the pre-test risk assessment and rationale, providing an understanding of how these factors contribute to the overall safety and usability of the system being tested. Our post-test risk is included and discussed in the results that follow.

Task #	*Task/Risk Level*	*Risk Rational*
1	User configures evidence-based DSI Moderate	Failure to configure evidence-based DSI properly could lead to inaccurate decision-making, affecting clinical outcomes.
2	User records source attributes for evidence-based DSI. Low	Minimal risk as it involves recording data elements already part of clinical workflows.
3	User changes source attributes for evidence-based DSI. Moderate	Changes to source attributes may affect the accuracy of clinical recommendations, leading to inappropriate care.
4	User accesses source attributes for evidence-based DSI. Moderate	Misinterpretation of source attributes could result in errors in clinical decision-making.
5	User selects Decision Support Intervention(s) based on any of the required elements Low	Selection based on predefined elements reduces the likelihood of user error.
6	Access source attributes for selected evidence-based DSI. Low	Accessing source attributes involves reviewing existing data, with a low likelihood of user error impacting clinical outcomes
7	Provide feedback for a triggered evidence-based DSI. Low	Feedback is non-intrusive and primarily involves confirming previously recorded actions, which limits the risk.
8	User exports feedback data in a computable format, including the data identified in (b)(11)(ii)(C) at a minimum. Low	Exporting data is a routine task, with minimal risk of affecting clinical outcomes.
9	Configures Predictive DSI using the required USCDI data elements.	Incorrect configuration could result in poor predictive outcomes, impacting patient care.

	Moderate	
10	User records user-defined source attributes for a Predictive DSI.	Low risk, as this task involves recording predefined data elements.
	Low	
11	User changes user-defined source attributes for a Predictive DSI.	Incorrect interpretation of user-defined attributes could lead to inaccuracies in the predictive model.
	Moderate	
12	User accesses user-defined source attributes for a Predictive DSI.	Low risk, since this is a basic access task with minimal potential for error.
	Low	
13	User selects a user-supplied Predictive DSI.	Selection errors could result in incorrect clinical predictions, affecting patient management.
	Moderate	
14	Access and reviews source attributes for selected user-supplied Predictive DSI.	Reviewing attributes carries minimal risk, as it typically involves verifying already recorded data.
	Low	

Results

The results of the usability test were calculated according to the methods specified in the Usability Metrics section. Participants who failed to follow session and task instructions had their data excluded from the analysis. There was no testing irregularities recorded.

The usability testing results for the EHRUT are detailed below. The results should be seen in light of the objectives and goals outlined in section on Study Design. The data should yield actionable results that, if corrected, yield material, positive impact on user performance.

The results from the Likert scale scored the subjective satisfaction with the system based on performance with these tasks to broadly interpreted. Scores under 3 represent poor usability and scores over 3 would be considered above average.

§170.315 (b)(11) Decision Support Intervention – Evidence Based DSI

Data Analysis and Reporting

Task #	Task	Scale	Task Rating	Task Rating - Std Dev.	Task Time - Mean(s)	Task Time - Standard Deviation(s)	Time - Observed/Optimal	Task Success - Mean (%)	Task Success - Std. Deviation(s)	Task Errors - Mean (%)	Task Error - Std. Deviation (%)	Observed - (# of Steps)	Optimal (# of Steps)
1	User configures evidence-based DSI using any of the required elements alone or in combination.	Likert	5	0	44.5	6.81	44/40	100	0	0	0	11	11
2	User records source attributes for evidence-based DSI.	Likert	5	0	26.6	4.39	27/22	100	0	0	0	3	3
3	User changes source attributes for evidence-based DSI.	Likert	5	0	57.9	8.43	57.5/50	100	0	0	0	6	6
4	User accesses source attributes for evidence-based DSI.	Likert	5	0	28.8	4.01	28.78/25	100	0	0	0	4	4
5	User selects Decision Support Intervention(s) based on any of the required elements alone or in combination.	Likert	4	.5	37	4.70	37/30	100	0	0	0	3	3

6	User accesses source attributes for selected evidence-based DSI.	Likert	5	0	43.7	5.27	43/35	100	0	0	0	3	3
7	User provides feedback for a triggered evidence-based DSI.	Likert	5	0	122.3	22.81	122/100	100	0	0	0	4	4
8	User exports feedback data in a computable format, including the data identified in (b)(11)(ii)(C) at a 5minimum.	Likert	5	0	56.6	10.06	56/40	100	0	0	0	3	3

Efficiency

Tasks in this group were generally completed efficiently, with users finding the interfaces intuitive. However, tasks that required detailed feedback (Task 19) or involved system-dependent actions (Task 20) occasionally led to delays. Minor interface inefficiencies, such as dropdown responsiveness and field navigation, were noted.

Effectiveness

All participants successfully completed the tasks (100% overall), demonstrating a clear understanding of objectives and processes. The intuitive design of most tasks supported error-free execution.

Satisfaction

Users expressed high levels of satisfaction, particularly for tasks with well-structured interfaces. Feedback highlighted simplicity and clarity as key strengths, though there were calls for improvements in system responsiveness and visual guidance.

Major findings

These tasks showed a consistent ability to meet objectives, with minor variability in task completion times. Tasks involving feedback or export functions revealed opportunities for optimization, especially in terms of system performance.

Post Test Risk Assessment and Remarks

Task #	*Task/Pre-test Risk Level*	*Test Error Percentage*	*Discussion*
1	User configures evidence-based DSI. Moderate	0%	No errors recorded. The configuration of the DSI was completed successfully, validating that users can accurately set up evidence-based interventions without issues.
2	User records source attributes for evidence-based DSI. Low	0%	Zero errors observed. Users effectively recorded source attributes, supporting the assumption that this task carries minimal risk when recording pre-defined data elements.
3	User changes source attributes for evidence-based DSI. Moderate	0%	No issues noted. The process of changing source attributes was done without error, demonstrating that changes can be made safely, maintaining clinical decision-making integrity.
4	User accesses source attributes for evidence-based DSI. Moderate	0%	No errors were encountered. The users successfully accessed source attributes, confirming the low likelihood of user misinterpretation or errors in clinical settings.
5	User selects Decision Support Intervention(s) based on any of the required elements. Low	0%	No errors observed. Selection of DSIs based on predefined elements was straightforward, reinforcing the minimal risk for user error during this task.
6	Access source attributes for selected evidence-based DSI. Low	0%	Task completed without errors. Users were able to access source attributes with ease, affirming that this is a low-risk task involving existing data.
7	Select DSI based on the problems, medications, allergies, and intolerances incorporated from a C-CDA. Low	0%	No errors recorded. Selection of DSIs based on C-CDA data went smoothly, indicating the system's ability to ensure accurate and up-to-date information from clinical documents.
8	Provide feedback for a triggered evidence-based DSI. Low	0%	Zero errors observed. Users were able to provide feedback without issues, confirming the task's low risk and the non-intrusive nature of this functionality.

Areas for improvement

Enhance system performance for data export (Task 20).

Streamline feedback forms with pre-filled fields or auto-completion options (Task 19).

Improve dropdown menu responsiveness and field labeling for easier navigation (Task 17).

Consider adding tooltips and quick-access features to simplify attribute selection and review processes (Tasks 13, 18).

§170.315 (b)(11) Decision Support Intervention – User-supplied Predictive DSI

Data Analysis and Reporting

Task #	Task	Scale	Task Rating	Task Rating - Std Dev.	Task Time - Mean(s)	Task Time - Standard Deviation(s)	Time - Observed/Optimal	Task Success - Mean (%)	Task Success - Std. Deviation(s)	Task Errors - Mean (%)	Task Error - Std. Deviation (%)	Observed - (# of Steps)	Optimal (# of Steps)
9	User configures Predictive DSI using the required USCDI data elements.	Likert	4	0	138.8	29.07	138/120	100	0	0	0	4	4
10	User records user-defined source attributes for a Predictive DSI.	Likert	5	.5	87.6	14.52	87/75	100	0	0	0	3	3
11	User changes user-defined source attributes for a Predictive DSI.	Likert	5	0	30.6	4.09	30/25	100	0	0	0	3	3
12	User accesses user-defined source attributes for a Predictive DSI.	Likert	5	0	70.7	10.27	70.74/60	100	0	0	0	3	3
13	User selects a user-supplied Predictive DSI.	Likert	5	.35	28.4	4.63	28.42/22	100	0	0	0	3	3
14	User accesses and reviews source attributes for selected user-supplied Predictive DSI.	Likert	5	0	80.5	14.03	84.47/70	100	0	0	0	3	3

Discussion of Findings

Efficiency

These tasks, particularly those requiring configuration or detailed review (Tasks 21, 26), were more time-consuming due to the complexity of predictive elements and detailed user-defined attributes. Tasks involving access and selection (Tasks 23, 25) were completed more quickly and consistently.

Effectiveness

All participants successfully completed these tasks (100% overall), though some required additional time for configuration and attribute changes. Tasks involving user-defined attributes showed a higher learning curve but were still effective.

Satisfaction

Users were generally satisfied with the clarity of instructions and the straightforward nature of most tasks. However, tasks with more complexity (Tasks 21, 26) received feedback suggesting the need for more interactive guidance or step-by-step instructions.

Major findings

The complexity of predictive DSI tasks led to longer completion times and more variability in user performance. Tasks related to accessing or modifying user-defined attributes were straightforward but could benefit from enhanced visual grouping.

Post Test Risk Assessment and Remark

Task #	*Task/Pre-test Risk Level*	*Test Error Percentage*	*Discussion*
9	Configures Predictive DSI using the required USCDI data elements. Moderate	0%	No errors were recorded. Configuration of the predictive DSI using USCDI data elements was successful, demonstrating that users can perform this moderately complex task without negatively impacting patient care.
10	User records user-defined source attributes for a Predictive DSI. Low	0%	Task completed without error. Users were able to record user-defined source attributes without issues, confirming the low risk associated with this task
11	User changes user-defined source attributes for a Predictive DSI. Moderate	0%	No errors observed. Accessing user-defined attributes was done smoothly, validating the system's ability to reduce the likelihood of misinterpretation during this process.
12	User accesses user-defined source attributes for a Predictive DSI. Low	0%	Zero errors. As expected, this basic task was completed without any challenges, supporting the minimal potential for error in this process.
13	User selects a user-supplied Predictive DSI. Moderate	0%	No errors were noted. Selection of a user-supplied Predictive DSI was performed correctly, minimizing the risk of incorrect clinical predictions affecting patient management.
14	Access and reviews source attributes for selected user-supplied Predictive DSI. Low	0%	No issues occurred. Users successfully reviewed source attributes, confirming the task's low risk as it typically involves verifying previously recorded

Areas for improvement

- Simplify the configuration process for predictive DSI by breaking it into smaller, guided steps (Task 21).
- Improve field labels and consider adding a search function to assist with attribute changes (Task 24).
- Provide visual summaries and highlight key attributes to streamline review processes (Task 26).
- Enhance grouping and contextual help for user-defined attributes (Task 22).

Appendices

Appendix A - Trademarks

Holy Name® is a registered trademark

Appendix B - Tasks

§170.315 (b)(11)- Decision Support Intervention – Evidence Based

<table>
<tr><th>Task No.</th><th>Description</th></tr>
<tr><td rowspan="6">1</td><td>Configure and enable Evidence-based DSI
Verify that users can configure an evidence-based DSI using any required elements such as problems, medications, allergies, intolerances, or any combination thereof.</td></tr>
<tr><td>Actor</td></tr>
<tr><td>Clinic Manager (Admin)</td></tr>
<tr><td>Steps</td></tr>
<tr><td>1. Start Login - Visit https://ehr.justtest.in/account/login.

2. Log in with the credentials:

• Username: (provided to test participant)

• Password: (provided to test participant)

3 Click 'Select Facility.'

4 In 'Patient Search,' enter 'Tom' in the 'First Name' field and click 'Search.'

5 Select 'Tom Harry' from the results.

6 Click 'Launch DSI App' (it will open in a new tab).

7. Enter the login credentials for the app:

• Username: provider

• Password: provider

8. Click 'Yes, Allow' on the next page.

9. Click 'Evidence Based Alerts' to start configuration of Evidence-based DSI for the patient.
10. Select DSI launch for combination of problems, labs and allergies.
11. Select “Evidence Based Alert’ to finish the task</td></tr>
<tr><td>Observations
<table>
<tr><th>Task Success</th><th>Path Deviations</th><th>Errors</th><th>Effort: (1) v. high, (5) v. low</th><th>Time to Complete</th></tr>
<tr><td>☒Pass ☐Fail</td><td>☒No ☐Yes</td><td>☒No ☐Yes</td><td>☐1 ☐2 ☐3 ☐4 ☒5</td><td>40 secs</td></tr>
</table></td></tr>
<tr><td>Comments
Click here</td></tr>
</table>

Task No.	Description
2	**User records source attributes for evidence-based DSI.** Confirm that users can record and store source attributes for evidence-based DSIs

Actor

Clinic Manager (Admin)

Steps

1. From current page select 'Evidence Based Alerts' and select 'Edit' navigate to the source attributes section.
2. Examine the required evidence-based source attributes (bibliographic citation, developer information, etc.).
3. 'Save' the record and verify the attributes are stored correctly.

Observations

Task Success	Path Deviations	Errors	Effort: (1) v. high, (5) v. low	Time to Complete
☒Pass ☐Fail	☒No ☐Yes	☒No ☐Yes	☐1 ☐2 ☐3 ☐4 ☒5	22 secs

Comments

Click here

Task No.	Description
3	**User changes source attributes for evidence-based DSI** Ensure users can modify the source attributes for a configured DSI.

Actor

Clinic Manager (Admin)

Steps

1. From the current page use the navigation "Back" function or arrow
2. From current page select 'Evidence Based Alerts' and select 'Edit' navigate to the source attributes section.
3. Examine the required evidence-based source attributes (bibliographic citation, developer information, etc.).
4. Modify the bibliographic citation by typing "JAMA" over the existing field
5. Modify the existing source attribute "revision date" to 2024.
6. Save changes on the bottom of the screen

Observations

Task Success	Path Deviations	Errors	Effort: (1) v. high, (5) v. low	Time to Complete
☒Pass ☐Fail	☒No ☐Yes	☒No ☐Yes	☐1 ☐2 ☐3 ☐4 ☒5	50 secs

Comments

Click here

Task No.	Description
4	**User accesses source attributes for evidence-based DSI** Verify that users can access the modified source attributes of an evidence-based DSI
	Actor
	Clinic Manager (Admin)
	Steps
	1. From the current page use the navigation “Back” function or arrow 2. From current page select ‘Evidence Based Alerts’ and select ‘Edit’ to navigate to the source attributes section. 3. Visually inspect the source attribute fields. 4. Confirm that all attributes are available for review and that Bibliographic Reference now says “JAMA” and the Revision Date says “2024”
	Observations

Task Success	Path Deviations	Errors	Effort: (1) v. high, (5) v. low	Time to Complete
☒Pass ☐Fail	☒No ☐Yes	☒No ☐Yes	☐1 ☐2 ☐3 ☐4 ☒5	25 secs

	Comments
	Click here

Task No.	Description
5	**User selects Decision Support Intervention(s) based on any of the required elements alone or in combination** Confirm that users can select DSIs based on a combination of required elements for problems, medications, and allergies.
	Actor
	Clinic User
	Steps
	1. Log in as an authorized user. 2. Select a DSI based on multiple required elements (e.g., problems + medications + allergies). 3. Activate the DSI and verify it triggers appropriately during patient interaction.
	Observations

Task Success	Path Deviations	Errors	Effort: (1) v. high, (5) v. low	Time to Complete
☒Pass ☐Fail	☒No ☐Yes	☒No ☐Yes	☐1 ☐2 ☐3 ☐4 ☒5	30 secs

	Comments
	Click here

Task No.	Description
6	**User accesses source attributes for selected evidence-based DSI** Ensure that source attributes for a selected evidence-based DSI are accessible.
	Actor
	Clinic Manager (Admin)
	Steps
	1. Select an active evidence-based DSI. 2. Navigate to the source attributes section. 3. Verify that the relevant source attributes are accessible and up to date. Review each field.
	Observations

Task Success	Path Deviations	Errors	Effort: (1) v. high, (5) v. low	Time to Complete
☒Pass ☐Fail	☒No ☐Yes	☒No ☐Yes	☐1 ☐2 ☐3 ☐4 ☒5	35 secs

Comments
Click here

Task No.	Description
7	**User provides feedback for a triggered evidence-based DSI** Ensure that users can provide feedback on a triggered DSI
	Actor
	Clinic User
	Steps
	1. Select “Evidence Based DSI” for any patient 2. Select “Evidence Based Alerts” 3. To the left of the respective alert provide feedback in the following fields: feedback, action, intervention, and remarks. 4. Ensure fields are populated and that text is “sticky”
	Observations

Task Success	Path Deviations	Errors	Effort: (1) v. high, (5) v. low	Time to Complete
☒Pass ☐Fail	☒No ☐Yes	☒No ☐Yes	☐1 ☐2 ☐3 ☐4 ☒5	100 secs

Comments
Click here

Task No.	Description
8	**Generate feedback in computable export with specific fields** Verify that feedback data can be exported with required fields and in a computable format
	Actor
	Clinic Manager (Admin)
	Steps
	1. From the current screen select “Export” for any of the alerts 2. Ensure the file for Feedback Export download commences in a computable format (.json) 3. Review the file for the following fields: user, date, location, action, intervention, and feedback/remarks
	Observations

Task Success	Path Deviations	Errors	Effort: (1) v. high, (5) v. low	Time to Complete
☒Pass ☐Fail	☒No ☐Yes	☒No ☐Yes	☐1 ☐2 ☐3 ☐4 ☒5	40 secs

Comments
Click here

§170.315 (b)(11)- Decision Support Intervention – User-supplied Predictive

Task No.	Description
9	**User configures Predictive DSI using the required USCDI data elements** Verify that users can configure predictive DSIs using USCDI data elements such as demographics, problems, and vital signs.
	Actor
	Clinic Manager (Admin)
	Steps
	Log in as a user with administrative rights. Navigate to the "Predictive DSI" section. Configure a predictive DSI using patient demographics, problems, and vital signs. Activate the DSI and verify that it uses the required USCDI data elements.
	Observations

Task Success	Path Deviations	Errors	Effort: (1) v. high, (5) v. low	Time to Complete
☒Pass ☐Fail	☒No ☐Yes	☒No ☐Yes	☐1 ☐2 ☐3 ☐4 ☒5	120 secs

Comments
Click here

Task No.	Description
10	**User records user-defined source attributes for a Predictive DSI** Ensure users can record custom source attributes for a predictive DSI.
	Actor
	Clinic Manager (Admin)
	Steps
	1. Select a predictive DSI and navigate to the source attributes section. 2. Record user-defined attributes, such as the intended use, developer details, and purpose of the DSI. 3. Save the attributes and confirm they are recorded correctly.
	Observations

Task Success	Path Deviations	Errors	Effort: (1) v. high, (5) v. low	Time to Complete
☒Pass ☐Fail	☒No ☐Yes	☒No ☐Yes	☐1 ☐2 ☐3 ☐4 ☒5	75 secs

	Comments
	Click here

Task No.	Description
11	**User changes user-defined source attributes for a Predictive DSI** Confirm that users can change the source attributes defined for a predictive DSI.
	Actor
	Clinic Manager (Admin)
	Steps
	1. Access a configured predictive DSI. 2. Navigate to the source attributes section and record a user-defined attributes. 3. Verify all attributes are visible and up to date based on the previous modification/edit.
	Observations

Task Success	Path Deviations	Errors	Effort: (1) v. high, (5) v. low	Time to Complete
☒Pass ☐Fail	☒No ☐Yes	☒No ☐Yes	☐1 ☐2 ☐3 ☐4 ☒5	25 secs

	Comments
	Click here

Task No.	Description
12	**User accesses user-defined source attributes for a Predictive DSI.** Confirm that users can access user defined source attributes defined for a predictive DSI.
	Actor
	Clinic Manager (Admin)
	Steps
	• Access a configured predictive DSI. • Navigate to the source attributes section and change 1 of the user-defined attributes. • Verify all attributes are visible and up to date.
	Observations

Task Success	Path Deviations	Errors	Effort: (1) v. high, (5) v. low	Time to Complete
☒Pass ☐Fail	☒No ☐Yes	☒No ☐Yes	☐1 ☐2 ☐3 ☐4 ☒5	60 secs

Comments
Click here

Task No.	Description
13	**User selects a user-supplied Predictive DSI.** Verify that users can select a predictive DSI configured with user-supplied attributes
	Actor
	Clinic User or Admin
	Steps
	1. Log in as a user with predictive DSI access. 2. Select a predictive DSI from the list of available interventions. 3. Confirm the DSI activates and generates recommendations based on user-supplied data.
	Observations

Task Success	Path Deviations	Errors	Effort: (1) v. high, (5) v. low	Time to Complete
☒Pass ☐Fail	☒No ☐Yes	☒No ☐Yes	☐1 ☐2 ☐3 ☐4 ☒5	22 secs

Comments
Click here

Task No.	Description				
14	**User accesses and reviews source attributes for selected user-supplied Predictive DSI.** Ensure that users can access and review source attributes for selected user-supplied predictive DSIs.				
	Actor				
	Clinic User				
	Steps				
	1. Select a user-supplied predictive DSI. 2. Access the source attributes related to the intervention. 3. Review the attributes (e.g., developer information, intended use) and confirm that they are accurate.				
	Observations				
	Task Success	**Path Deviations**	**Errors**	**Effort: (1) v. high, (5) v. low**	**Time to Complete**
	☒Pass ☐Fail	☒No ☐Yes	☒No ☐Yes	☐1 ☐2 ☐3 ☐4 ☒5	70 secs
	Comments				
	Click here				

Appendix C - Consent to Remote Testing

Consent Form: Remote Usability Test

Please read and sign this form.

During this usability test I agree to participate in an online session using my computer and telephone. During the session I will be interviewed about the site, asked to find information or complete tasks using the site and asked to complete an online questionnaire about the experience.

I understand and consent to the use and release of the recording by . I understand that the information and recording are for research purposes only and that my name and image will not be used for any other purpose. I relinquish any rights to the recording and understand the recording may be copied and used by without further permission.

I understand that participation is voluntary, and I agree to immediately raise any concerns you might have.

If you have any questions after today, please contact <developer name>

Please sign below to indicate that you have read and understand the information on this form and that any questions you might have about the session have been answered.

Please print your name:

_____

Please sign your name:

_____

Participant's Signature or eSignature

Today's Date:_____

Thank you!

We appreciate your participation.

Test: ___*/ /* _____ *to* _____*/ /*____